



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 111157

TO: Nirmal Basi
Location: cm1/10e17
Art Unit: 1646
Friday, January 02, 2004

Case Serial Number: 09/719485

From: Mary Jane Ruhl
Location: Biotech-Chem Library
CM1-6A06
Phone: 605-1155

maryjane.ruhl@uspto.gov

Search Notes

Examiner Basi,

Here are the results for your recent search request.

Please feel free to contact me if you have any questions about these results.

Thank you for using STIC services. We appreciate the opportunity to serve you.

Sincerely,

Mary Jane Ruhl
Technical Information Specialist
STIC
CM-1, Rm. 6-A-06
605-1155

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 31, 2003, 18:55:58 ; Search time 4719.89 Seconds
(without alignments)
10739.032 Million cell updates/sec

Title: US-09-719-485-2

Perfect score: 1239

Sequence: 1 atggcgagccctggaacgg.....acgtgaagacatgggataa 1239

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_on.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sta.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_on.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sta.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_mam.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1239	100.0	1239	6	AX154589 Sequence
2	1239	100.0	1239	6	AX549187 Sequence
3	1239	100.0	1239	6	AX572965 Sequence
4	901	72.7	1161	6	AX154591 Sequence
5	901	72.7	2040	9	AF034632 Homo sapi
6	901	72.7	163284	9	AL137000 Human DNA
7	901	72.7	341560	2	AL596304 Homo sapi
8	901	72.7	349980	6	AX711879 Sequence
9	901	72.7	349980	6	AX739961 Sequence
10	592.2	47.8	692	9	HS3339407 Homo sapi
11	572.2	46.2	692	9	HS3339408 Homo sapi
12	557.4	45.0	615	9	HS3339459 Homo sapi
13	424.8	34.3	701	9	HS3326768 Homo sapi
14	307	24.8	192116	2	AC109158 Mus muscu
15	283	22.8	283	6	AR168467 Sequence
16	283	22.8	283	6	AR182284 Sequence
17	283	22.8	283	6	AR303910 Sequence
18	283	22.8	283	6	BD056681 Galanin r
19	283	22.8	283	6	BD064652 Human gal
20	283	22.8	283	6	BD064664 Mouse gal
21	266.8	21.5	573	9	HS3341295 Homo sapi
22	244	19.7	1751	5	AB095995 Gallus ga
23	242.6	19.6	1351	5	AB095997 Gallus ga
24	242.6	19.6	1703	5	AB095996 Gallus ga
25	242.6	19.6	4121	5	AB095994 Gallus ga
26	234.8	19.0	1063	6	AR156351 Sequence
27	234.8	19.0	1101	4	SSU80178 Sus scrofa
28	233.2	18.8	870	4	SSU60180 Sus scrofa
29	230	18.6	1029	6	AR156352 Sequence
30	230	18.6	110000	2	AC095360 2 Continuation (3 of
31	230	18.6	183099	2	AC117349 Rattus no
32	230	18.6	195131	2	AC108400 Mus muscu
33	230	18.6	227035	2	AC116585 Mus muscu
34	230	18.6	228796	2	AC095173 Rattus no
35	230	18.6	250754	2	AC130771 Rattus no
36	229.8	18.5	1092	6	AR156357 Sequence
37	229.8	18.5	1092	6	BD094703 GHSR liga
38	229.8	18.5	1092	6	ES4510 Novel poly
39	229.8	18.5	1350	10	AB001982 Rattus no
40	229	18.5	250	6	AR269796 Sequence
41	228.8	18.5	870	6	AX548854 Sequence
42	228.8	18.5	870	9	HSU60181 Human growt
43	228.8	18.5	1122	6	AR156354 Sequence
44	228.8	18.5	6787	9	AF369786 Homo sapi
45	228.8	18.5	145593	9	AC069523 Homo sapi

ALIGNMENTS

RESULT 1	AX154589	AX154589	1239 bp	DNA	linear	PAT 22-JUN-2001
LOCUS	Sequence 9 from Patent WO0138355.					
DEFINITION	AX154589					
ACCESSION	AX154589					
VERSION	AX154589.1	GI:14536175				
KEYWORDS						
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1					
AUTHORS	Sheppard,P.O., Jaspers,S.R., Deisher,T.A. and Bishop,P.D.					
TITLE	Method of forming a peptide-receptor complex with zaig33 and therapeutic use thereof					


```

Qy 541 CAGGACCCCGCATCTCCGTAGTCCCGGGCTCAATGGCACCGCGGATCGCTCTCG 600
Db 541 CAGGACCCCGCATCTCCGTAGTCCCGGGCTCAATGGCACCGCGGATCGCTCTCG 600
Qy 601 CCTCTCGCTCGTCCCGCTCTCTCGTCTCGCGGGGCGCACCGCGTCCCGCGCTCG 660
Db 601 CCTCTCGCTCGTCCCGCTCTCTCGTCTCGCGGGGCGCACCGCGTCCCGCGCTCG 660
Qy 661 GGGCCGAGACCGCGGAGCGCGGCTGTTACGCGCGGAAATCGCGCGGAGCGCGCG 720
Db 661 GGGCCGAGACCGCGGAGCGCGGCTGTTACGCGCGGAAATCGCGCGGAGCGCGCG 720
Qy 721 CAGTGGCGCGCTGCGTGTGTCATCTGTGGGTCAACCGCTACTTCTTCTCCCTTT 780
Db 721 CAGTGGCGCGCTGCGTGTGTCATCTGTGGGTCAACCGCTACTTCTTCTCCCTTT 780
Qy 781 CTGTGCTCAGCATCTCTACGGCTCATCGGGCTCATCGGGCGGAGCTGTGGAGCAGCGCGCG 840
Db 781 CTGTGCTCAGCATCTCTACGGCTCATCGGGCTCATCGGGCGGAGCTGTGGAGCAGCGCGCG 840
Qy 841 CTGGAGCGCGCGCTCGGGCGGAGAGGCCACCGGACAGCGTCCGGTCTTG 900
Db 841 CTGGAGCGCGCGCTCGGGCGGAGAGGCCACCGGACAGCGTCCGGTCTTG 900
Qy 901 CTGTGCTGCTTCTGCGCATTTAATTTGCTGTTGCTTCCACGTTGGCAGATCAT 960
Db 901 CTGTGCTGCTTCTGCGCATTTAATTTGCTGTTGCTTCCACGTTGGCAGATCAT 960
Qy 961 TACATAACGAGAAATTCGGGATGATGATCTTCTCAGTACTTTTAAATCATCGTCT 1020
Db 961 TACATAACGAGAAATTCGGGATGATGATCTTCTCAGTACTTTTAAATCATCGTCT 1020
Qy 1021 CTGCAACTTTTCTATCTGAGCGCATCTCAACCCATCTCTACACCTCATTTCAA 1080
Db 1021 CTGCAACTTTTCTATCTGAGCGCATCTCAACCCATCTCTACACCTCATTTCAA 1080
Qy 1081 AAGTACAGAGCGCGCTTTAACTGCTGCTCAAGAGAGTCCAGCGCGAGAGGCTTC 1140
Db 1081 AAGTACAGAGCGCGCTTTAACTGCTGCTCAAGAGAGTCCAGCGCGAGAGGCTTC 1140
Qy 1141 CACAGAAGCGGACACTCGCGGGGAAAGTTGCGAGGAGACACTGGAGGAGACACGGTGGG 1200
Db 1141 CACAGAAGCGGACACTCGCGGGGAAAGTTGCGAGGAGACACTGGAGGAGACACGGTGGG 1200
Qy 1201 TACACCGAGACAGCGCTAAGCTGAAGCATGGGATA 1239
Db 1201 TACACCGAGACAGCGCTAAGCTGAAGCATGGGATA 1239

```

```

RESULT 4
AX154591
LOCUS AX154591 1161 bp DNA linear PAT 23-JUN-2001
DEFINITION Sequence 11 from Patent WO0138355.
ACCESSION AX154591
VERSION AX154591.1 GI:14536177
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Sheppard,P.O., Jaspers,S.R., Delisher,T.A. and Bishop,P.D.
TITLE Method of forming a peptide-receptor complex with zsig33 and
therapeutic use thereof
JOURNAL Patent: WO 0138355-A 11 31-MAY-2001;
ZymoGenetics, Inc. (US)
FEATURES
source
1. .1161
/organism="Homo sapiens"
/db_type="genomic DNA"
/db_xref="taxon:9606"
1. .1161
/note="unnamed protein product"
CDS

```

```

/codon_start=1
/protein_id="CAC42612.1"
/db_xref="GI:14536178"
/translation="MSPWNGSDPEGAREPPHPALPPCDERRCSPPFLGALVPTAV
CLCLFVVGSGNVVTVMLIGRYDMRTTNYLYGSMVADLLILLGLGFPFDIYLRWR
PWFGPLLCRLSLYVGEGETYATLLHMTALSVRYLAICRPLRARVLVTRRRVALLA
VLWAVALLGAAFLFLVGPQDFGI SVVPGNGTARIASPLASSPPLMFSRAPPPSP
PSPETAEEAALFSECRPSPAOLGALRWMLVTTTAVFFLPFLCLSLYGLIGRELWS
SRPLRGPASGRGHRGTQVRLRWKSRGSKDACLQAPPGTAOTGLPLPLLAQLW
APLAPPPISIPASTRRGGSGIYNLLVALPRWQNLHLKHGRFADDDVLLSL"
BASE COUNT 126 a 436 c 372 g 227 t
ORIGIN
Query Match 72.7% Score 901; DB 6; Length 1161;
Best Local Similarity 100.0%; Pred. No. 1.5e-110; Indels 0; Gaps 0;
Matches 901; Conservative 0; Mismatches 0;
Qy 1 ATGGGAGCCCTTGGAAACGCGAGCGAGCGCCCGAGGGGGCGGGAGCGCCGTGGCCC 60
Db 1 ATGGGAGCCCTTGGAAACGCGAGCGAGCGCCCGAGGGGGCGGGAGCGCCGTGGCCC 60
Qy 61 GCGCTGCGCGCTTGGAGAGAGCGCGCTGCTCGCCCTTTCCCTTGGGGGGCGCTGGTGG 120
Db 61 GCGCTGCGCGCTTGGAGAGAGCGCGCTGCTCGCCCTTTCCCTTGGGGGGCGCTGGTGG 120
Qy 121 GTGACCGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 121 GTGACCGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Qy 181 ATGCTGATCGGGGGCTTACCGGGAGCATGCGGACCAACCACTTTGTACCTGGGCGAG 240
Db 181 ATGCTGATCGGGGGCTTACCGGGAGCATGCGGACCAACCACTTTGTACCTGGGCGAG 240
Qy 241 GCGGTGTCCGACCTACTCATCTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 300
Db 241 GCGGTGTCCGACCTACTCATCTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 300
Qy 301 TCGCGGCGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Db 301 TCGCGGCGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Qy 361 TGACCTTACGCCACGCTGTGTCATGACCGCGCTCAGCGCTCAGCGCTACCTGGCCATC 420
Db 361 TGACCTTACGCCACGCTGTGTCATGACCGCGCTCAGCGCTCAGCGCTACCTGGCCATC 420
Qy 421 TCGCGGCGCTGCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Db 421 TCGCGGCGCTGCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Qy 481 GTGCTGTGGGCGTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Db 481 GTGCTGTGGGCGTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Qy 541 CAGGACCCCGGCATCTCCGTAGTCCCGGGCTCAATGGCACCGCGGATCGCTCTCTCG 600
Db 541 CAGGACCCCGGCATCTCCGTAGTCCCGGGCTCAATGGCACCGCGGATCGCTCTCTCG 600
Qy 601 CTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 601 CTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Qy 661 GGGCCGAGACCGCGGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 661 GGGCCGAGACCGCGGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Qy 721 CAGCTGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 721 CAGCTGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Qy 781 CTGTGCTCAGCATCTCTACGGCTCATCGGGCTCATCGGGCGGAGCTGTGGAGCAGCG 840
Db 781 CTGTGCTCAGCATCTCTACGGCTCATCGGGCTCATCGGGCGGAGCTGTGGAGCAGCG 840

```

```
Qy 841 CTGCGAGCCCGCGCCCTCGGGCGGAGAGAGCCACCGGACAGCCGTCCGCGTCTG 900
Db 841 CTGCGAGCCCGCGCCCTCGGGCGGAGAGAGCCACCGGACAGCCGTCCGCGTCTG 900
Qy 901 C 901
Db 901 C 901

RESULT 5
AF034632 2040 bp DNA linear PRI 13-JUL-1998
LOCUS Homo sapiens orphan G protein-coupled receptor (GPR38) gene,
DEFINITION complete cds.
ACCESSION AF034632
VERSION AF034632.1 GI:2654158
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2040)
AUTHORS McKee,K.K., Tan,C.P., Palyha,O.C., Liu,J., Feighner,S.D.,
Hreniuk,D.L., Smith,R.G., Howard,A.D. and Van der Ploeg,I.H.
TITLE Cloning and characterization of two human G protein-coupled
receptor genes (GPR38 and GPR39) related to the growth hormone
secretagogue and neurotensin receptors
JOURNAL Genomics 46 (3), 426-434 (1997)
MEDLINE 98110578
PUBMED 9441746
REFERENCE 2 (bases 1 to 2040)
AUTHORS McKee,K.K., Tan,C.P., Palyha,O.C., Liu,J., Feighner,S.D.,
Hreniuk,D.L., Smith,R.G., Van Der Ploeg,I.H.T. and Howard,A.D.
TITLE Direct Submission
JOURNAL Submitted (17-NOV-1997) Biochemistry and Physiology, Merck and Co.,
Inc., PO Box 2000, Rahway, NJ 07065, USA
FEATURES
source
1. .2040
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="13"
/map="13q14-21"
<1..>2040
/gene="GPR38"
join(<1..901,1703..>2040)
/gene="GPR38"
join(1..901,1703..2040)
/note="orphan G protein-coupled receptor related to growth
hormone secretagogue and neurotensin receptors"
/codon_start=1
/product="GPR38"
/protein_id="AAC26081.1"
/db_xref="GI:2654159"
/translation="MGSPWNGSDGPEAREPPWALPCDERRCSPPPLGALVPVAV
CLFLFVGVSGNVVTLIGRYDMRTTNLYLGSMAVSDLLILGLPFDLYRLWRSR
PWFGPLLCLRLSLYGECTYATLLHMTALSYERYLAICRPLRLRVLTERRVRLAIA
VIMAVALLSAGFFLFLVGEQDPGISVVGUNGRTARIASSPLASPLWLSRAPPPSP
PSGPETAAALFLFRECPSPQAQLGRLVMTVTAFFLPLCLSLYLGLIGRELWS
SRRLRGPAASGRGRHQTVRLVLLVLAFLIICWLPFHVGRIIVINTEDSRMYFSQ
YFNIVALQLFVLSAINPILNLYLSKRYRAAFAKLLARKSRPRGHRSDRTAGEVAG
DTGDDTVGYTETSANVTMG"
BASE COUNT 347 a 631 c 598 g 464 t
ORIGIN
Query Match 72.7%; Score 901; DB 9; Length 2040;
Best Local Similarity 100.0%; Pred. No. 1.3e-110;
Matches 901; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGCAGCCCTTGAAACGGAGCGACCGGCCCGAGGGCGCGGAGCCCGCGTGGCCC 60
Db 1 ATGGCAGCCCTTGAAACGGAGCGACCGGCCCGAGGGCGCGGAGCCCGCGTGGCCC 60
```

```
Qy 61 GCGCTGCGCGCTTTCGACGAGCGCGCTGCTCGCCCTTTCCCTTGGGGGGCGCTGGTCCG 120
Db 61 GCGCTGCGCGCTTTCGACGAGCGCGCTGCTCGCCCTTTCCCTTGGGGGGCGCTGGTCCG 120
Qy 121 GTGACCGCTGTGTGCTGTGCTGCTGTTGCTGTGGGGTGAAGCGCAACGTGGTACCCTG 180
Db 121 GTGACCGCTGTGTGCTGTGCTGTTGCTGTGGGGTGAAGCGCAACGTGGTACCCTG 180
Qy 181 ATGCTGATCGGGCGCTACCGGGACATCGGACACCAACCTTGTACCTGGGCGAGCATG 240
Db 181 ATGCTGATCGGGCGCTACCGGGACATCGGACACCAACCTTGTACCTGGGCGAGCATG 240
Qy 241 GCGGTGTCGACCTACTCATCTGCTCGGGCTGCGCTTTCGACCTGTATCCGCTCTGCGC 300
Db 241 GCGGTGTCGACCTACTCATCTGCTCGGGCTGCGCTTTCGACCTGTATCCGCTCTGCGC 300
Qy 301 TCGGGCGCTTGGTGTTCGGGGCGCTGCTTTCGGCGCTGCTCCCTTACGTGGGGGAGGGC 360
Db 301 TCGGGCGCTTGGTGTTCGGGGCGCTGCTTTCGGCGCTGCTCCCTTACGTGGGGGAGGGC 360
Qy 361 TGCACTTACGCCAGCTGTGCACATGACCGCGCTCAGCGTTCGAGCGCTACCTGGCCATC 420
Db 361 TGCACTTACGCCAGCTGTGCACATGACCGCGCTCAGCGTTCGAGCGCTACCTGGCCATC 420
Qy 421 TGGCGCCCGCTCGCGCGCGCGCTCTTGGTCAACCGCGCGCGCTTCGCGCGCTCATCGCT 480
Db 421 TGGCGCCCGCTCGCGCGCGCGCTCTTGGTCAACCGCGCGCGCTTCGCGCGCTCATCGCT 480
Qy 481 GTGCTTGGGGCGGTGGCGCTGCTCTTGGCGGCTCCCTTCTTGTTCCTGGGGCGTGGAG 540
Db 481 GTGCTTGGGGCGGTGGCGCTGCTCTTGGCGGCTCCCTTCTTGTTCCTGGGGCGTGGAG 540
Qy 541 CAGGACCCCGCATCTCCGTAGTCCCGGGCTCAATGGCACCGCGCGGATCGCCCTCTCG 600
Db 541 CAGGACCCCGCATCTCCGTAGTCCCGGGCTCAATGGCACCGCGCGGATCGCCCTCTCG 600
Qy 601 CCTCTCGCTCTGTCGCGCGCTCTCTGCTCTCGGGCGCGCACCGCGCTTCCCGCGCTCG 660
Db 601 CCTCTCGCTCTGTCGCGCGCTCTCTGCTCTCGGGCGCGCACCGCGCTTCCCGCGCTCG 660
Qy 661 GGGCCCGAGACCGCGAGCGCGCGCTGTTTACGCGCGCAATGTCGGCGGAGCCCGCG 720
Db 661 GGGCCCGAGACCGCGAGCGCGCGCTGTTTACGCGCGCAATGTCGGCGGAGCCCGCG 720
Qy 721 CAGCTGGGGCGGTGCGCTGTATGCTGTGGGTACACCGCGCTACTTCTTCTGCGCTTT 780
Db 721 CAGCTGGGGCGGTGCGCTGTATGCTGTGGGTACACCGCGCTACTTCTTCTGCGCTTT 780
Qy 781 CTGTGCTTACGATCTCTACGGGCTCATCGGGCGGAGCTGTGGAGCAGCGCGCGCG 840
Db 781 CTGTGCTTACGATCTCTACGGGCTCATCGGGCGGAGCTGTGGAGCAGCGCGCGCG 840
Qy 841 CTGCGAGCGCGCGCGCTCGGGCGGAGAGAGCCACCGGACAGCCGTCCGCGTCTG 900
Db 841 CTGCGAGCGCGCGCGCTCGGGCGGAGAGAGCCACCGGACAGCCGTCCGCGTCTG 900
Qy 901 C 901
Db 901 C 901

RESULT 6
AL137000 163284 bp DNA linear PRI 18-DEC-2000
LOCUS Human DNA sequence from clone RP11-203116 on chromosome 13 Contains
DEFINITION the gene for KIAA0970 protein, COX7CP1 (cytochrome c oxidase
subunit VIIC pseudogene 1), a novel pseudogene, the GPR38 (G
protein-coupled receptor 38) gene, ESTs, STSs, GSSs and a CpG
island, complete sequence.
ACCESSION AL137000
VERSION AL137000.6 GI:19944121
KEYWORDS HTG; COX7CP1; CpG island; cytochrome c oxidase; G protein-coupled
```

receptor; GPR38; KIAA0970.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 163284)
 REFERENCE Direct Submission
 AUTHORS Submitted (18-DEC-2000) Sanger Centre, Hinxton, Cambridgeshire,
 TITLE CB10 15A, UK. E-mail enquiries: humquery@sanger.ac.uk
 JOURNAL requests: clonerequest@sanger.ac.uk
 COMMENT On Aug 29, 2000 this sequence version replaced gi:9926419.
 During sequence assembly data is compared from overlapping clones.
 where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 The following abbreviations are used to associate primary accession
 numbers given in the feature table with their source databases:
 Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information
 on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
 was generated from part of bacterial clone contigs of human
 chromosome 13, constructed by the Sanger Centre Chromosome 13
 Mapping Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr13>
 IMPORTANT: This sequence is not the entire insert of clone
 RP11-203116. It may be shorter because we sequence overlapping
 sections only once, except for a 100 base overlap.
 The true left end of clone RP11-452110 is at 1 in this sequence.
 The true right end of clone RP11-103J18 is at 163284 in this
 sequence. This sequence has been finished according to sequence map
 criteria as follows. An attempt is made to resolve all sequencing
 problems, such as compressions and repeats, but not necessarily
 within known annotated repeat sequence elements. Where the
 sequence is ambiguous, there is an annotation using the 'unsure'
 feature key. RP11-203116 is from the library RPCI-11.1 constructed
 by the group of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
 VECTOR: pBAC3.6.
 Location/Qualifiers
 1..163284
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="13"
 /clone="RP11-203116"
 /clone_lib="RPCI-11.1"
 3..121
 /note="MER20 repeat: matches 4..126 of consensus"
 repeat_region 152..261
 /note="MIR repeat: matches 20..137 of consensus"
 repeat_region 262..554
 /note="AluX repeat: matches 1..293 of consensus"
 repeat_region 555..698
 /note="MIR repeat: matches 137..255 of consensus"
 repeat_region 1012..1291
 /note="AluX repeat: matches 7..291 of consensus"
 repeat_region 1320..1425
 /note="MER24 repeat: matches 4..114 of consensus"
 repeat_region 1591..1804
 /note="L1P8 repeat: matches 5945..6158 of consensus"
 repeat_region 3240..3621
 /note="L2 repeat: matches 2310..2691 of consensus"
 repeat_region 3824..4106
 /note="AluX repeat: matches 1..301 of consensus"
 repeat_region 4975..5042
 /note="L2 repeat: matches 2639..2705 of consensus"
 repeat_region 6113..6408
 /note="AluX repeat: matches 1..295 of consensus"
 repeat_region 7646..7732
 /note="AluSc repeat: matches 1..85 of consensus"
 9887..10191
 /note="AluX repeat: matches 1..306 of consensus"
 repeat_region 12201..12247
 /note="L2 repeat: matches 2704..2750 of consensus"
 repeat_region 12472..12738
 /note="L2 repeat: matches 1304..1530 of consensus"
 repeat_region 12739..12874
 /note="FLAM C repeat: matches 2..133 of consensus"
 repeat_region 12875..12887
 /note="L2 repeat: matches 1530..1542 of consensus"
 repeat_region 12922..13227
 /note="L1MB3 repeat: matches 5871..6183 of consensus"
 repeat_region 13303..13348
 /note="MTR1E repeat: matches 1..46 of consensus"
 repeat_region 13533..13623
 /note="MTR1E repeat: matches 223..315 of consensus"
 repeat_region 13857..13865
 /note="MTR1E repeat: matches 307..533 of consensus"
 repeat_region 13890..14150
 /note="MTR1J repeat: matches 195..456 of consensus"
 repeat_region 14151..14450
 /note="AluX repeat: matches 1..307 of consensus"
 repeat_region 14501..14729
 /note="L1MB repeat: matches 588..813 of consensus"
 repeat_region 14730..15006
 /note="AluX repeat: matches 1..285 of consensus"
 repeat_region 15007..15361
 /note="L1MB repeat: matches 813..1178 of consensus"
 repeat_region 15389..15437
 /note="L1MB repeat: matches 1159..1207 of consensus"
 repeat_region 15438..16412
 /note="HERV repeat: matches 2891..3879 of consensus"
 repeat_region 16413..16616
 /note="MTR2B repeat: matches 1..216 of consensus"
 repeat_region 17092..17141
 /note="MER5B repeat: matches 27..79 of consensus"
 repeat_region 17416..17647
 /note="MIR repeat: matches 17..261 of consensus"
 repeat_region 19402..19499
 /note="HY3 repeat: matches 2..99 of consensus"
 repeat_region 20730..21163
 /note="L2 repeat: matches 2168..2609 of consensus"
 repeat_region 25800..25825
 /note="13 copies 2 mer tt 92% conserved"
 repeat_region 26041..26131
 /note="FLAM C repeat: matches 17..107 of consensus"
 repeat_region 26144..26249
 /note="53 copies 2 mer tt 64% conserved"
 repeat_region 27029..27259
 /note="MIR repeat: matches 2..251 of consensus"
 repeat_region 27422..27788
 /note="L2 repeat: matches 1879..2259 of consensus"
 repeat_region 27863..28239
 /note="L2 repeat: matches 1278..1734 of consensus"
 repeat_region 28255..28571
 /note="AluY repeat: matches 1..310 of consensus"
 repeat_region 28729..29128
 /note="L1MB3 repeat: matches 4905..5326 of consensus"
 repeat_region 29170..29413
 /note="L1MB3 repeat: matches 5461..5715 of consensus"
 repeat_region 29448..29681
 /note="L2 repeat: matches 1594..1828 of consensus"
 repeat_region 29683..29886
 /note="L1MB repeat: matches 5527..5734 of consensus"
 repeat_region 29888..30323
 /note="L2 repeat: matches 2129..2596 of consensus"
 repeat_region 30324..30620
 /note="AluX repeat: matches 1..299 of consensus"
 repeat_region 30621..30780
 /note="L2 repeat: matches 2596..2749 of consensus"
 repeat_region 30784..30813
 /note="15 copies 2 mer aa 86% conserved"
 repeat_region 30814..30947


```
|||||
38856 CTGCGAGCCCGCGCGCTCGGGCGGAGAGAGCCACCGGACCGTCCGGTCTG 38915
QY 901 C 901
Db 38916 C 38916

RESULT 10
HSA339407 HSA339407 692 bp DNA linear PRI 18-JUL-2002
LOCUS Homo sapiens genomic sequence surrounding NotI site, clone
DEFINITION NR5-1B14C.
ACCESSION AJ339407
VERSION AJ339407.1 GI:15883825
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 692)
AUTHORS Kutsenko,A.S., Gizatullin,R.Z., Al-Amin,A.N., Wang,F., Kvasha,S.M.,
Podowski,R.M., Matushkin,Y.G., Gyanchandani,A., Muravenko,O.V.,
Levitsky,V.G., Kolchanov,N.A., Protodopov,A.I., Kashuba,V.I.,
Kisselev,L.L., Wasserman,W., Wahlestedt,C. and Zabarovsky,E.R.
TITLE NotI flanking sequences: a tool for gene discovery and verification
of the human genome
JOURNAL Nucleic Acids Res. 30 (14), 3163-3170 (2002)
MEDLINE 22131767
PUBMED 12136098
REFERENCE 2 (bases 1 to 692)
AUTHORS Zabarovsky,E.R.
TITLE Direct Submission
JOURNAL Submitted (16-MAY-2001) Microbiology and Tumorbiology Centre,
Karolinska Institute, Theorells vag, 3, Box 280, Stockholm 171 77,
Sweden
FEATURES
source Location/Qualifiers
1..692
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="NR5-1B14C"
BASE COUNT 61 a 256 c 236 g 129 t 10 others
ORIGIN
Query Match 47.8%; Score 592.2; DB 9; Length 692;
Best Local Similarity 92.9%; Pred. No. 1.7e-69;
Matches 627; Conservative 0; Mismatches 43; Indels 5; Gaps 1;
QY 1 ATGGGAGCCCTGGAAACGGCAGCGCGCCCGAGGGGGCGGGAGCGCGTGGCCC 60
Db 20 ATGGGAGCCCTGGAAACGGCAGCGCGCCCGAGGGGGCGGGAGCGCGTGGCCC 79
QY 61 GCGTGGCGGCTTGCAGACGAGCGCGTGTCTGCGCCCTTTCCCTGGGGGGCGTGGTCCG 120
Db 80 GCGTGGCGGCTTGCAGACGAGCGCGTGTCTGCGCCCTTTCCCTGGGGGGCGTGGTCCG 139
QY 121 GTGACCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 180
Db 140 GTGACCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 199
QY 181 ATGCTGATCGGGCGCTACCGGACATCGGACACACCACTGTACCTGGGCGAGCATG 240
Db 200 ATGCTGATCGGGCGCTACCGGACATCGGACACACCACTGTACCTGGGCGAGCATG 259
QY 241 GCGGTGTCGACCTACTCATCTCTCTGCGGCTGCGGCTTGCACCTGTACCGCCTCTGCGCG 300
Db 260 GCGGTGTCGACCTACTCATCTCTCTGCGGCTGCGGCTTGCACCTGTACCGCCTCTGCGCG 319
QY 301 TCGGGCGGCTGTGGTTCGGGCGGCTGTCTGCGGCGGCTGTCTCTCTCTCTCTCTCTCTCTCT 360
Db 320 TCGGGCGGCTGTGGTTCGGGCGGCTGTCTGCGGCGGCTGTCTCTCTCTCTCTCTCTCTCTCT 379
```

```
361 TGACCTACGCCACGCTGTGTGACATGACCGCGCTCAGCGCTCAGCGCTACCTGGCCATC 420
Db 380 TGACCTACGCCACGCTGTGTGACATGACCGCGCTCAGCGCTCAGCGCTACCTGGCCATN 439
QY 421 TGCGGCGGCTCGCGCGCGCGCTTGGTCAACCGGCGCGCGCTCCGCGGCGCTCATCGCT 480
Db 440 TGCGGCGGCTCGCGCGCGCGCTTGGTCAACCGGCGCGCGCTCCGCGGCGCTCATCGCT 499
QY 481 GTGCTGTGGGCGGTGGGCGGTCTCTGCGGCGTCCCTTCTTGTCTGTGGGCGGTGGAG 540
Db 500 GTGCTGTGGGCGGTGGGCGGTCTCTGCGGCGTCCCTTCTTGTCTGTGGGCGGTGGAG 559
QY 541 CAGGACCCCGGCATCTCCGTAGTCCCGGGCTCAATGGCACCGCGCGGATCGCTCCCTCG 600
Db 560 CAGGACCCCGGCATCTCCGTAGTCCCGGGCTCAATGGCACCGCGGATCGCTCCCTCG 619
QY 601 CTCTCGCTGTGCGCGCGCTCTCTGCGGCGCGCACCGCGCTCCCGCGCGCTCG 660
Db 620 CTCTTGTGCTGTGCGCGCGCTCTCTGCGGCGCGCACCGCGCTCCCGCGCGCTCG 674
QY 661 GGGCGCGGAGACCGCG 675
Db 675 GGGCGCGGAGACCGCG 689

RESULT 11
HSA342408 HSA342408 692 bp DNA linear PRI 18-JUL-2002
LOCUS Homo sapiens genomic sequence surrounding NotI site, clone
DEFINITION NR1-PB24C.
ACCESSION AJ342408
VERSION AJ342408.1 GI:15886883
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 692)
AUTHORS Kutsenko,A.S., Gizatullin,R.Z., Al-Amin,A.N., Wang,F., Kvasha,S.M.,
Podowski,R.M., Matushkin,Y.G., Gyanchandani,A., Muravenko,O.V.,
Levitsky,V.G., Kolchanov,N.A., Protodopov,A.I., Kashuba,V.I.,
Kisselev,L.L., Wasserman,W., Wahlestedt,C. and Zabarovsky,E.R.
TITLE NotI flanking sequences: a tool for gene discovery and verification
of the human genome
JOURNAL Nucleic Acids Res. 30 (14), 3163-3170 (2002)
MEDLINE 22131767
PUBMED 12136098
REFERENCE 2 (bases 1 to 692)
AUTHORS Zabarovsky,E.R.
TITLE Direct Submission
JOURNAL Submitted (16-MAY-2001) Microbiology and Tumorbiology Centre,
Karolinska Institute, Theorells vag, 3, Box 280, Stockholm 171 77,
Sweden
FEATURES
source Location/Qualifiers
1..692
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="NR1-PB24C"
BASE COUNT 92 a 239 c 219 g 140 t 2 others
ORIGIN
Query Match 46.2%; Score 572.2; DB 9; Length 692;
Best Local Similarity 91.5%; Pred. No. 7.7e-67;
Matches 616; Conservative 0; Mismatches 55; Indels 2; Gaps 1;
QY 1 ATGGGAGCCCTGGAAACGGCAGCGCGCCCGAGGGGGCGGGAGCGCGTGGCCC 60
Db 20 ATGGGAGCCCTGGAAACGGCAGCGCGCCCGAGGGGGCGGGAGCGCGTGGCCC 79
QY 61 GCGTGGCGGCTTGCAGACGAGCGCGTGTCTGCGCCCTTTCCCTGGGGGGCGTGGTCCG 120
Db 80 GCGTGGCGGCTTGCAGACGAGCGCGTGTCTGCGCCCTTTCCCTGGGGGGCGTGGTCCG 139
```


Qy 833 GCGGCGCGCTGCGAGGCGCGCCCTCGGGCGG 867
|||
Db 92236 GGGTCCCGCACCAAGGTGCGAAGAAAGGGCGG 92202
|||

RESULT 15

AR168467
LOCUS AR168467 283 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 4 from patent US 6287855.

ACCESSION AR168467

VERSION AR168467.1 GI:17904400

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE Unclassified.

1 (bases 1 to 283)

AUTHORS Tan, C. and Sullivan, K.

TITLE Nucleic acid encoding rat galanin receptor (GALR2)

JOURNAL Patent: US 6287855-A 4 11-SEP-2001;

FEATURES Location/Qualifiers

source 1..283

BASE COUNT 27 a 116 c 84 g 56 t

ORIGIN

Query Match 22.8%; Score 283; DB 6; Length 283;

Best Local Similarity 100.0%; Pred. No. 2.5e-28;

Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 206 TCGGACACACCAACTTGTACCTGGGACGATGGCGGTGTCGACCTACTATCCTGC 265

Db 1 TCGGACACACCAACTTGTACCTGGGACGATGGCGGTGTCGACCTACTATCCTGC 60

Qy 266 TCGGGCTGCCGTTCCGACCTGTACCGCCTCTGGCGCTCGCGCCCTGGGTTCGGGCGG 325

Db 61 TCGGGCTGCCGTTCCGACCTGTACCGCCTCTGGCGCTCGCGCCCTGGGTTCGGGCGG 120

Qy 326 TGCTCTGCCGCTGTCCCTCTACGTGGGCGAGGGCTGCACCTAGCCACGCTGCTGCACA 385

Db 121 TGCTCTGCCGCTGTCCCTCTACGTGGGCGAGGGCTGCACCTAGCCACGCTGCTGCACA 180

Qy 386 TGACCGGCTCAGCGTCGAGCGTACCTGGCCATCTCGCGCCGCTCCGCGCCGCGTCT 445

Db 181 TGACCGGCTCAGCGTCGAGCGTACCTGGCCATCTCGCGCCGCTCCGCGCCGCGTCT 240

Qy 446 TGGTCACCGCGCGCGGTCCGCGGCTCATCGCTGTGCTCTG 488

Db 241 TGGTCACCGCGCGCGGTCCGCGGCTCATCGCTGTGCTCTG 283

Search completed: January 1, 2004, 01:09:43

Job time : 4728.39 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 31, 2003, 13:48:11 ; Search time 380.324 Seconds
(without alignments)
8794.080 Million cell updates/sec

Title: US-09-719-485-2

Perfect score: 1239

Sequence: 1 atggcgagcccttgaacgg.....acgtgaagacatgggataa 1239

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_19Jun03.*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
- 25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1239	100.0	1239	21	AAZ45403
2	1239	100.0	1239	22	AAF83683
3	1239	100.0	1239	22	AAF85449
4	1239	100.0	1239	24	ABK90132
5	1239	100.0	1239	25	ABZ42842
6	1239	100.0	1506	22	AAI66989
7	1078	87.0	1390	21	AAZ45404
8	915.6	73.9	1203	22	AAF85448

9	901	72.7	1161	22	AAF83684	Short form of motl
10	901	72.7	3066	21	AAZ45402	Genomic sequence o
11	896.2	72.3	2040	21	AAA46116	Human G protein co
12	546	44.1	813	22	AAF85447	Nucleotide sequenc
13	501.2	40.5	1179	24	ABQ47146	Oligonucleotide fo
14	501.2	40.5	1179	24	ABQ47147	Oligonucleotide fo
15	437.2	35.3	1179	24	ABQ47148	Oligonucleotide fo
16	437.2	35.3	1179	24	ABQ47149	Oligonucleotide fo
17	283	22.8	283	19	AAV4930	Galanin receptor G
18	283	22.8	283	19	AAV32651	Galanin receptor G
19	283	22.8	283	19	AAV28390	Galanin receptor G
20	283	22.8	283	19	ABK14060	Rat galanin recept
21	264.8	21.4	1092	21	AAZ45405	DNA encoding the p
22	264.8	21.4	1092	22	AAF85450	Nucleotide sequenc
23	255.8	20.6	1050	21	AAZ61492	cDNA encoding cani
24	234.8	19.0	1063	18	AAZ68662	Pig growth hormone
25	234.8	19.0	1063	18	AAZ69754	Pig growth hormone
26	233.2	18.8	1029	18	AAZ68663	Pig growth hormone
27	233.2	18.8	1029	18	AAZ69755	Swine growth hormo
28	231.6	18.7	1095	21	AAZ45993	cDNA encoding the
29	231.6	18.7	4009	21	AAZ45967	DNA encoding the m
30	229.8	18.5	1092	22	AAZ27800	Rat growth hormone
31	229	18.5	250	25	ACA55761	Pig signalling pat
32	228.8	18.5	870	25	ABZ42674	Human growth hormo
33	228.8	18.5	1122	18	AAZ68665	Human growth hormo
34	228.4	18.4	3129	18	AAZ68667	Rat growth hormone
35	228.4	18.4	3129	18	AAZ69759	Rat growth hormone
36	227.8	18.4	1088	18	AAZ68664	Human growth hormo
37	227.8	18.4	1088	18	AAZ69756	Human growth hormo
38	227.8	18.4	1101	21	AAA30643	Human G protein-co
39	227.8	18.4	1101	21	AAA30732	DNA encoding human
40	227.8	18.4	1101	21	AAZ51463	Human G protein-co
41	227.8	18.4	1101	21	AAZ83680	Human G protein co
42	227.8	18.4	1101	24	AAZ30395	Human growth-hormo
43	227.8	18.4	1101	25	ACA56550	Human signalling p
44	227.2	18.3	1122	18	AAZ69757	Human growth hormo
45	226.8	18.3	1092	18	AAZ69760	Rat growth hormone

ALIGNMENTS

RESULT 1
AAZ45403
ID AAZ45403 standard; cDNA; 1239 BP.

XX AAZ45403;

AC AAZ45403;

XX 27-MAR-2000 (first entry)

XX cDNA encoding the motilin receptor splice variant MTL-R1A.

XX Motilin receptor; G-protein coupled receptor; MTL-R1; GPR38;
XX spliced form; MTL-R1a; MTL-R1b; gastric motility disorder;
XX functional defect; neurological disorder; scleroderma; colonoscopy;
XX paraneoplastic syndrome; radiation induced dysmotility; diabetes;
XX infection; stress-related motility disorder; psychogenic disorder;
XX gastroparesis; gastro-esophageal reflux disease; constipation;
XX chronic idiopathic pseudo obstruction; acute faecal impaction;
XX postoperative ileus; gallstones; infantile colic; diarrhoea;
XX irritable bowel syndrome; non-ulcer dyspepsia; non-cardiac chest pain;
XX endoscopy; duodenal intubation; ds.

OS Homo sapiens.

XX Key Location/Qualifiers

XX CDS 1..1239

XX FT /*tag= a

XX FT /product= "MTL-R1A"

XX WO9964436-A1.

XX PD 16-DEC-1999.

XX Human; G protein-coupled receptor 38; receptor; GPR38; gene; ds;
KW Alzheimer's disease; Parkinson's disease; ulcerative colitis;
KW Crohn's disease; Hodgkin's disease; glioblastoma; breast carcinoma;
KW colon carcinoma; lung small cell carcinoma; lung adenocarcinoma;
KW pancreatic small cell carcinoma; pancreatic adenocarcinoma.
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT CDS 1..1239
FT /*tag= a
FT /product= "Human G protein-coupled receptor 38 (GPR38)"
XX
PN WO200257791-A2.
PD 25-JUL-2002.
XX
PF 29-NOV-2001; 2001WO-US45219.
PR 29-NOV-2000; 2000US-250251P.
PR 30-NOV-2000; 2000US-250452P.
XX
PA (LIFE-) LIFESPAN BIOSCIENCES INC.
XX
PI Brown JP, Burmer GC, Roush CL, Kulander BG;
XX WPI; 2002-566812/60.
DR P-PSDB; ABG30936.
XX
PT Assay for detecting Alzheimer's disease, Parkinson's disease,
PT ulcerative colitis, Crohn's disease, Hodgkin's disease, glioblastoma or
PT carcinoma, comprises using a binding partner for G protein coupled
PT receptor 38 -
XX
PS Disclosure; Fig 1; 112pp; English.
CC
CC The present invention relates to a new assay method that involves
CC contacting a binding partner specific for G protein coupled receptor
CC (GPR) 38 with specific cells. The method of the invention is useful for
CC the detection of an increased risk of Alzheimer's disease, Parkinson's
CC disease, ulcerative colitis, Crohn's disease, Hodgkin's disease,
CC glioblastoma, or carcinoma. GPR 38 is used to manufacture a medicament
CC for inhibiting, treating or preventing Alzheimer's disease, Parkinson's
CC disease, ulcerative colitis, Crohn's disease, Hodgkin's disease,
CC glioblastoma, breast carcinoma, colon carcinoma, lung small cell
CC carcinoma, lung adenocarcinoma, pancreatic small cell carcinoma, and
CC pancreatic adenocarcinoma. An agonist or antagonist to GPR 38 are used
CC to manufacture a medicament able to reduce the symptoms of these
CC diseases. Nucleic acids encoding GPR 38 can also be used to treat the
CC diseases. The present nucleic acid sequence encodes the human G protein-
CC coupled receptor 38 (GPR38) of the invention.
XX
SQ Sequence 1239 BP; 172 A; 431 C; 397 G; 239 T; 0 other;
Query Match 100.0%; Score 1239; DB 24; Length 1239;
Best Local Similarity 100.0%; Pred. No. 1.9e-216;
Matches 1239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGAGAGCCCTGGAAACGGCGACGACGGCCCGCGAGGGGCGCGAGCGCGCGTGGCCC 60
DB 1 ATGGGAGAGCCCTGGAAACGGCGACGACGGCCCGCGAGGGGCGCGAGCGCGCGTGGCCC 60
QY 61 GGGCTGCCCTTGGACAGCGCCGCTGCTCGCCCTTTCCCTTGGGGGCGTGGTGGCG 120
DB 61 GGGCTGCCCTTGGACAGCGCCGCTGCTCGCCCTTTCCCTTGGGGGCGTGGTGGCG 120
QY 121 GTGACCGCTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 180
DB 121 GTGACCGCTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 180
QY 181 ATGCTGATCGGGCGCTACCGGGAGCATGCGGACCAACCACTTGTACCTGGGCGAGCATG 240
DB 181 ATGCTGATCGGGCGCTACCGGGAGCATGCGGACCAACCACTTGTACCTGGGCGAGCATG 240

181 ATGCTGATCGGGCGCTACCGGGAGCATGCGGACCAACCACTTGTACCTGGGCGAGCATG 240
QY 241 GCGGTGTCGACCTACTACTCTCTGCTCGGGTGCCTTTCGACCTGTACCGCTCTGGGCGC 300
DB 241 GCGGTGTCGACCTACTACTCTCTGCTCGGGTGCCTTTCGACCTGTACCGCTCTGGGCGC 300
QY 301 TCGCGCCCTGGGTGTTTGGGGCGCTGCTCTGCGCGCTGTCTCTCTCTCTCTCTCTCTCTCT 360
DB 301 TCGCGCCCTGGGTGTTTGGGGCGCTGCTCTGCGCGCTGTCTCTCTCTCTCTCTCTCTCTCT 360
QY 361 TGCACCTAGCCACGCTGCTGCACATGACCGGCTCAGCGCTACCGCTTACCTGGGCGATC 420
DB 361 TGCACCTAGCCACGCTGCTGCACATGACCGGCTCAGCGCTACCGCTTACCTGGGCGATC 420
QY 421 TCGCGCCCGCTCGCGCGCGCTCTTGGTTCACCCCGCGCGCTCGCGCGCTCATCGCT 480
DB 421 TCGCGCCCGCTCGCGCGCGCTCTTGGTTCACCCCGCGCGCTCGCGCGCTCATCGCT 480
QY 481 GTGCTCTGGGCGCTGGGCGCTGCTCTCTGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
DB 481 GTGCTCTGGGCGCTGGGCGCTGCTCTCTGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
QY 541 CAGGACCCCGCGCATCTCGGTAGTCCCGGCTCAATGGCACCGCGGATCGCTCTCTCTCT 600
DB 541 CAGGACCCCGCGCATCTCGGTAGTCCCGGCTCAATGGCACCGCGGATCGCTCTCTCTCT 600
QY 601 CCTCTCGCTCGTCCGCGCTCTCTGCGCTCTCGCGGCGCACCGCGCTCGCGCGCTCG 660
DB 601 CCTCTCGCTCGTCCGCGCTCTCTGCGCTCTCGCGGCGCACCGCGCTCGCGCGCTCG 660
QY 661 GGGCCCGAGACCGCGAGCGCGCGCTGTTTACCGCGGAATGCGGCGGAGCGCGCGCG 720
DB 661 GGGCCCGAGACCGCGAGCGCGCGCTGTTTACCGCGGAATGCGGCGGAGCGCGCGCGCG 720
QY 721 CAGCTGGGCGCGCTCGCTGCTCATGCTGGGTTCACACCGCTACTTCTCTCGGCTTT 780
DB 721 CAGCTGGGCGCGCTCGCTGCTCATGCTGGGTTCACACCGCTACTTCTCTCGGCTTT 780
QY 781 CTGTGCTCAGCATCTCTTACGGGCTCATCGGGCGGAGCTGTGAGCAGCGCGCGCG 840
DB 781 CTGTGCTCAGCATCTCTTACGGGCTCATCGGGCGGAGCTGTGAGCAGCGCGCGCGCG 840
QY 841 CTGCGAGCGCGCGCGCTCGGGCGGAGAGAGCGCACCGCGAGACCGCTCGCGCTCTG 900
DB 841 CTGCGAGCGCGCGCGCTCGGGCGGAGAGAGCGCACCGCGAGACCGCTCGCGCTCTG 900
QY 901 CTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
DB 901 CTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
QY 961 TACATAAACACGGAAGATTCGCGGATGATGATGATGATGATGATGATGATGATGATGATG 1020
DB 961 TACATAAACACGGAAGATTCGCGGATGATGATGATGATGATGATGATGATGATGATGATG 1020
QY 1021 CTGCAACTTTTCTATCTGAGCGCATCTATCAACCAATCTCTTACAACCTCATTTCAAAG 1080
DB 1021 CTGCAACTTTTCTATCTGAGCGCATCTATCAACCAATCTCTTACAACCTCATTTCAAAG 1080
QY 1081 AAGTACAGAGCGCGCGCTTTAAACTGCTGCTCGCAAGGAAGTCCAGGCGGAGAGCTTC 1140
DB 1081 AAGTACAGAGCGCGCGCTTTAAACTGCTGCTCGCAAGGAAGTCCAGGCGGAGAGCTTC 1140
QY 1141 CACAGAAGCAGGACACACGCGGGAAGTTGAGGGGACACCTGGAGGAGACACGCTGGGC 1200
DB 1141 CACAGAAGCAGGACACACGCGGGAAGTTGAGGGGACACCTGGAGGAGACACGCTGGGC 1200
QY 1201 TACACCGAGACAAAGCGCTAAACGCTGAAGACGATGGGATAA 1239
DB 1201 TACACCGAGACAAAGCGCTAAACGCTGAAGACGATGGGATAA 1239
RESULT 5
AB242842

Db 1021 CTGCAACTTTTCTATCTGAGCGCATCTATCAACCCCAATCTCTCAACCTCATTTCAAG 1080
Qy 1081 AAGTACAGAGCGCGCTTTAACTGCTGCTCGCAAGGAAGTCCAGGCGGAGAGCTTC 1140
Db 1081 AAGTACAGAGCGCGCTTTAACTGCTGCTCGCAAGGAAGTCCAGGCGGAGAGCTTC 1140
Qy 1141 CACAGAAGCAGGACACTCGGGGGGAAGTTGCAGGGGACACTGAGGAGACACGGTGGGC 1200
Db 1141 CACAGAAGCAGGACACTCGGGGGGAAGTTGCAGGGGACACTGAGGAGACACGGTGGGC 1200
Qy 1201 TACACCGAGACAGCGCTTAACGTGAAGACGATGGGATAA 1239
Db 1201 TACACCGAGACAGCGCTTAACGTGAAGACGATGGGATAA 1239

RESULT 6

AAI66989

ID AAI66989 standard; DNA; 1506 BP.

XX AAI66989;

AC AAI66989;

XX

DT 30-JAN-2002 (first entry)

DE Human GPR38 variant GPR38V polypeptide encoding DNA.

XX

XX GPR38V; variant; antibacterial; cytostatic; analgesic; antiasthmatic;

KW anti-Parkinsonian; hypertensive; hypotensive; antidiabetic; osteopathic;

KW antiallergic; antimigraine; neuroleptic; nootropic; anticonvulsant;

KW antiulcer; antiemetic; cardiac; vaccine; ds.

XX

OS Homo sapiens.

XX

FH Key

FT Location/Qualifiers

CDS 1..1506

FT /*tag= a

FT /product= "GPR38V"

FT

XX

PN W0200164836-A2.

XX

XX 07-SEP-2001.

XX

XX 28-FEB-2001; 2001WO-US06277.

XX

XX 01-MAR-2000; 2000US-0516315.

XX

XX (SMIK) SMITHKLINE BEECHAM CORP.

XX

XX Elshourbagy N, Shabon U;

XX

PI WPI; 2001-638956/73.

XX

DR P-PSDB; AAG65822.

XX

XX New human GPR38V polypeptide and polynucleotide, useful for treating

PT e.g. bacterial, fungal, protozoal and viral infections, cancers or

PT allergies, as vaccines, and for identifying agonists and antagonists

PT potentially useful in therapy -

XX

PS Claim 2; Page 26; 32pp; English.

XX

XX This DNA encodes a human GPR38 variant (GPR38V) polypeptide. GPR38V can

CC be expressed by standard recombinant methodology. The polynucleotides and

CC polypeptides are used in the treatment of bacterial, fungal, protozoal

CC and viral infections, infections caused by HIV-1 or HIV-2, pain, cancers,

CC diabetes, obesity, anorexia, asthma, Parkinson's disease, acute heart

CC failure, hypertension, urinary retentions, osteoporosis, allergies,

CC ulcers, migraine, psychotic and neurological disorders, or dyskinesias.

CC They are also useful for identifying agonists and antagonists that are

CC potentially useful in therapy, as vaccines to induce immunological

CC response in a mammal. The polypeptides may also be used as immunogens to

CC produce antibodies immunospecific for the polypeptides, and to identify

CC membrane bound or soluble receptors.

XX

XX Sequence 1506 BP; 207 A; 534 C; 509 G; 256 T; 0 other;

SQ

Query Match 100.0%; Score 1239; DB 22; Length 1506;
Best Local Similarity 100.0%; Pred. No. 1.9e-216;
Matches 1239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGGCGAGCCCTGGAAACGGCAGCGCGCCCGAGGGGGCGGAGCGCGCGTGGCCCC 60
Db 268 ATGGGCGAGCCCTGGAAACGGCAGCGCGCCCGAGGGGGCGGAGCGCGCGTGGCCCC 327
Qy 61 GGGCTGCCCGCTTGGCAGCAGCGCGCGCTGCTGCGCCCTTTCCCTGGGGGGCGTGGTGGCG 120
Db 328 GGGCTGCCCGCTTGGCAGCAGCGCGCGCTGCTGCGCCCTTTCCCTGGGGGGCGTGGTGGCG 387
Qy 121 GTGACCGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 388 GTGACCGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 447
Qy 181 ATGCTGATCGGGGGCTTACCGGGGACATGCGGACACCAACCACTTGTACCTGGGACGATG 240
Db 448 ATGCTGATCGGGGGCTTACCGGGGACATGCGGACACCAACCACTTGTACCTGGGACGATG 507
Qy 241 GCGGTGTCGAGACCTACTCATCTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 300
Db 508 GCGGTGTCGAGACCTACTCATCTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 567
Qy 301 TCGCGCCCTGGGTGTTTCGGGCGCGTGTCTGCGGCTGTCTCTCTCTCTCTCTCTCTCTCT 360
Db 568 TCGCGCCCTGGGTGTTTCGGGCGCGTGTCTGCGGCTGTCTCTCTCTCTCTCTCTCTCTCT 627
Qy 361 TGCACCTACGCCACCGCTGTGCACATGACCGGCGCTCAGGTCGAGCGCTACCTGGGCGATC 420
Db 628 TGCACCTACGCCACCGCTGTGCACATGACCGGCGCTCAGGTCGAGCGCTACCTGGGCGATC 687
Qy 421 TGCAGCGCGCTCGCGCGCGCGCTGCTTGGTCAACCGCGCGCGCTCGCGCGCGCTCATCGCT 480
Db 688 TGCAGCGCGCTCGCGCGCGCGCTGCTTGGTCAACCGCGCGCGCTCGCGCGCGCTCATCGCT 747
Qy 481 GTGCTCTGGGCGCGTGGCGCTGCTCTCTGCGGCTGCTCTCTCTCTCTCTCTCTCTCTCT 540
Db 748 GTGCTCTGGGCGCGTGGCGCTGCTCTCTGCGGCTGCTCTCTCTCTCTCTCTCTCTCTCT 807
Qy 541 CAGGACCCCGGCGATCTCCGTAGTCCCGGCGCTCAATGGCACCAGCGGCGGATCGGCTCTCG 600
Db 808 CAGGACCCCGGCGATCTCCGTAGTCCCGGCGCTCAATGGCACCAGCGGCGGATCGGCTCTCG 867
Qy 601 CTTCTCGCTGCTGCGCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 868 CTTCTCGCTGCTGCGCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 927
Qy 661 GGGCGCGAGACCGCGGAGCGCGCGCTGTTTTCAGCGCGGAAATGCGGCGGAGCGCGGCGCG 720
Db 928 GGGCGCGAGACCGCGGAGCGCGCGCTGTTTTCAGCGCGGAAATGCGGCGGAGCGCGGCGCG 987
Qy 721 CAGTGGGCGCGCTGCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 988 CAGTGGGCGCGCTGCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1047
Qy 781 CTGTGCTCAGCATCTCTTACGGGCTCATCGGGCGGAGCTGTGGAGCAGCGCGGCGCGCG 840
Db 1048 CTGTGCTCAGCATCTCTTACGGGCTCATCGGGCGGAGCTGTGGAGCAGCGCGGCGCGCG 1107
Qy 841 CTGCGAGGCGCGCGCGCTCGGGCGGAGAGAGGCGCACCGGCGAGACCGCTCGCGCTGCTG 900
Db 1108 CTGCGAGGCGCGCGCGCTCGGGCGGAGAGAGGCGCACCGGCGAGACCGCTCGCGCTGCTG 1167
Qy 901 CTGGTGGTGTCTGCGATTTAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Db 1168 CTGGTGGTGTCTGCGATTTAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1227
Qy 961 TACATAAACACGGAAGATTCGGGATGATGATCTCTCTCAGTACTTCTTAAACATGCTGCT 1020
Db 1228 TACATAAACACGGAAGATTCGGGATGATGATCTCTCTCAGTACTTCTTAAACATGCTGCT 1287

1021 CTGCACTTTTCTATCTGAGCGCATCTATCAACCCATCTCTACACCTCATTTCAAAG 1080
 1288 CTGCACTTTTCTATCTGAGCGCATCTATCAACCCATCTCTACACCTCATTTCAAAG 1347
 1081 AAGTACAGAGCGGGGCTTTAACTGCTGCTGCAAGAGTCCAGGCGGAGAGGCTTC 1140
 1348 AAGTACAGAGCGGGGCTTTAACTGCTGCTGCAAGAGTCCAGGCGGAGAGGCTTC 1407
 1141 CACAGAGAGGAGGACACTGCGGGGGAAGTTGCAAGGGGACACTGGAGGAGACAGGTGGGC 1200
 1408 CACAGAGAGGAGGACACTGCGGGGGAAGTTGCAAGGGGACACTGGAGGAGACAGGTGGGC 1467
 1201 TACACCGAGAGCAAGCGCTAACTGTAAGACGATGGATAA 1239
 1468 TACACCGAGAGCAAGCGCTAACTGTAAGACGATGGATAA 1506

RESULT 7
 AAZ45404
 ID AAZ45404 standard; cDNA; 1390 BP.
 XX
 AC AAZ45404;
 XX
 DT 27-MAR-2000 (first entry)
 XX
 DE cDNA encoding the motilin receptor splice variant MTL-R1B.
 XX
 KW Motilin receptor; G-protein coupled receptor; MTL-R1; GPR38;
 KW spliced form; MTL-R1A; MTL-R1B; gastric motility disorder;
 KW functional defect; neurological disorder; scleroderma; colonoscopy;
 KW paraneoplastic syndrome; radiation induced dysmotility; diabetes;
 KW infection; stress-related motility disorder; psychogenic disorder;
 KW gastroparesis; gastro-oesophageal reflux disease; constipation;
 KW chronic idiopathic pseudo obstruction; acute faecal impaction;
 KW postoperative ileus; gallstones; infantile colic; diarrhoea;
 KW irritable bowel syndrome; non-ulcer dyspepsia; non-cardiac chest pain;
 KW endoscopy; duodenal intubation; ds.
 XX
 OS Homo sapiens.

Key Location/Qualifiers
 CDS 1..1161
 FT /*tag= a
 FT /product= "MTL-R1B"
 XX
 PN W09964436-A1.
 XX
 PD 16-DEC-1999.
 XX
 XX 08-JUN-1999; 99WO-US12773.
 XX
 XX 12-JUN-1998; 98US-0089098.
 PR (MERI) MERCK & CO INC.
 XX
 XX Feighner SD, Patchett AA, Tan C, McKee K, MacNeil D, Howard AD;
 PI Pong S, Smith RG;
 PI
 XX WPI; 2000-105868/09.
 XX P-PSDB; AAY54146.
 XX
 PT Novel receptor protein for screening compounds used in treating
 PT irritable bowel syndrome, constipation and other gastric conditions
 XX
 PS Claim 6; Fig 4; 44pp; English.
 XX
 CC The present sequence encodes splice variant MTL-R1B of the motilin
 CC receptor. The gene encodes a G-protein coupled receptor, and is
 CC designated MTL-R1 (also GPR38). Two spliced forms of GPR38 exist,
 CC MTL-R1A (see AAY54145) and MTL-R1B (see AAY54146). MTL-R1A is a
 CC functional seven transmembrane domain form, and MTL-R1B is a truncated
 CC five transmembrane domain. The MTL-R1 proteins are used to identify
 CC agonists and antagonists which can be used for treating gastric motility

disorders, functional defects, disorders secondary to neurological
 disorders e.g. scleroderma, paraneoplastic syndromes radiation induced
 dysmotility, diabetes, infections, stress-related motility disorders,
 psychogenic disorders, gastroparesis, gastro-oesophageal reflux disease,
 constipation, chronic idiopathic pseudo obstruction, acute faecal
 impaction, postoperative ileus, gallstones, infantile colic, irritable
 bowel syndrome, non-ulcer dyspepsia, non-cardiac chest pain and
 diarrhoea. They can also be used in the preparation for colonoscopy,
 endoscopy and duodenal intubation.
 XX
 SQ Sequence 1390 BP; 190 A; 496 C; 436 G; 268 T; 0 other;
 Query Match 87.0%; Score 1078; DB 21; Length 1390;
 Best Local Similarity 89.1%; Pred. No. 3.6e-187;
 Matches 1239; Conservative 0; Mismatches 0; Indels 151; Gaps 1;
 QY 1 ATGGGAGAGCCCTTGGAAACGGACGAGCGGCGGCGGAGGGGCGGAGCGCGCTGGCCCC 60
 DB 1 ATGGGAGAGCCCTTGGAAACGGACGAGCGGCGGCGGAGGGGCGGAGCGCGCTGGCCCC 60
 QY 61 GCGCTGCGCGCTTGGACGAGCGCGCTGCTCGCCCTTTCCCTGGGGGCGTGGTCCCG 120
 DB 61 GCGCTGCGCGCTTGGACGAGCGCGCTGCTCGCCCTTTCCCTGGGGGCGTGGTCCCG 120
 QY 121 GTGACCGCTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 180
 DB 121 GTGACCGCTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 180
 QY 181 ATGCTGATCGGGCGCTACCGGACATGCGGACACACCAACCTTGTACCTGGGCGAGCATG 240
 DB 181 ATGCTGATCGGGCGCTACCGGACATGCGGACACACCAACCTTGTACCTGGGCGAGCATG 240
 QY 241 GCGCTGTCGACCTACTCATCTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 300
 DB 241 GCGCTGTCGACCTACTCATCTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 300
 QY 301 TCGCGGCGCTGCGGCTGCTGCGGCGCTGCTGCGGCGCTGCTGCGGCGCTGCTGCGGCGCTGCTG 360
 DB 301 TCGCGGCGCTGCGGCTGCTGCGGCGCTGCTGCGGCGCTGCTGCGGCGCTGCTGCGGCGCTGCTG 360
 QY 361 TGCACCTAGCCACCGTGTGTCATGACCGCGCTCAGCGCTCAGCGCTCAGCGCTCAGCGCTCAGCG 420
 DB 361 TGCACCTAGCCACCGTGTGTCATGACCGCGCTCAGCGCTCAGCGCTCAGCGCTCAGCGCTCAGCG 420
 QY 421 TGCAGCGCGCTGCGGCGCGCTGCTGCTGCGGCGCTGCTGCTGCGGCGCTGCTGCTGCGGCGCTGCTG 480
 DB 421 TGCAGCGCGCTGCGGCGCGCTGCTGCTGCGGCGCTGCTGCTGCGGCGCTGCTGCTGCGGCGCTGCTG 480
 QY 481 GTGCTCTGGGCGCTGCGGCGCTGCTGCTGCGGCGCTGCTGCTGCGGCGCTGCTGCTGCGGCGCTGCTG 540
 DB 481 GTGCTCTGGGCGCTGCGGCGCTGCTGCTGCGGCGCTGCTGCTGCGGCGCTGCTGCTGCGGCGCTGCTG 540
 QY 541 CAGGACCCCGGCGATCTCCGAGTCCCGGCGCTCAATGGCACCGCGCGGATGCGCTCTCTCG 600
 DB 541 CAGGACCCCGGCGATCTCCGAGTCCCGGCGCTCAATGGCACCGCGCGGATGCGCTCTCTCG 600
 QY 601 CCTCTGCGCTGCTGCGGCGCTGCTGCTGCGGCGCGCACCGCGCTGCGGCGCGCTGCGGCGCGCTGCG 660
 DB 601 CCTCTGCGCTGCTGCGGCGCTGCTGCTGCGGCGCGCACCGCGCTGCGGCGCGCTGCGGCGCGCTGCG 660
 QY 661 GGGCCCGAGACCGCGGAGGCGGCGCTGCTGAGCGCGGAAATGCGCGGCGAGCGCGCGCGCG 720
 DB 661 GGGCCCGAGACCGCGGAGGCGGCGCTGCTGAGCGCGGAAATGCGCGGCGAGCGCGCGCGCGCG 720
 QY 721 CAGCTGGGCGCGCTGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
 DB 721 CAGCTGGGCGCGCTGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
 QY 781 CTGTGCTCAGCATCTCTACGGGCTCATCGGCGGAGCTGTGGAGCAGCGCGCGCGCGCG 840
 DB 781 CTGTGCTCAGCATCTCTACGGGCTCATCGGCGGAGCTGTGGAGCAGCGCGCGCGCGCG 840
 QY 841 CTGCGAGGCGCGCGCTGCGGCGGAGAGAGGCGCACCGGCGAGACCGTCCGCGTCTCTG 900

Db 841 CTGCGAGCCCGCGCTCGGGCGGAGAGAGCCACCGGACACCGTCCGCGCTCTG 900
 Qy 901 C----- 901
 Db 901 CGTAAGTGGAGCCCGCGTGGTCCAAAGACGCCCTGCCTGCAGTCCGCCCGCGGGGACC 960
 Qy 902 ----- 901
 Db 961 GCGCAACGCTGGGTCCCTTCCCTGCTGCCCGAGCTCTGGGGCGGCTTCCAGCTCCC 1020
 Qy 902 -----TGGTGGTGGTCTGGCATTTATAATTG 929
 Db 1021 TTTCCTATTTCGATTCAGGCTCCACCGCGTGGTGGTCTGGCATTTATAATTG 1080
 Qy 930 CTGGTTCGCCCTCCACGCTGGCAGAACTATTACATAAAACAGGAAGATTGGCGGATGAT 989
 Db 1081 CTGGTGGCCCTCCACGCTGGCAGAACTATTACATAAAACAGGAAGATTGGCGGATGAT 1140
 Qy 990 GTACTTCTCAGPACTTTTAACATCGCTGCTGCTGCAACTTTTCTATCTGAGCGCATCTAT 1049
 Db 1141 GTACTTCTCAGPACTTTTAACATCGCTGCTGCTGCAACTTTTCTATCTGAGCGCATCTAT 1200
 Qy 1050 CAACCCCAATCTCTACAACTCATTTCAAGAGATCAGAGCGCGGCTTTAACTGCT 1109
 Db 1201 CAACCCCAATCTCTACAACTCATTTCAAGAGATCAGAGCGCGGCTTTAACTGCT 1260
 Qy 1110 GCTCGCAAGGAGTCCAGGCGGAGAGGCTTCCACAGAGGAGGAGACACTGCGGGGAAGT 1169
 Db 1261 GCTCGCAAGGAGTCCAGGCGGAGAGGCTTCCACAGAGGAGGAGACACTGCGGGGAAGT 1320
 Qy 1170 TGCAGGGGACACTGGAGGAGACAGCGTGGGCTACACCGAGACAGCGCTAACGTTGAAGAC 1229
 Db 1321 TGCAGGGGACACTGGAGGAGACAGCGTGGGCTACACCGAGACAGCGCTAACGTTGAAGAC 1380
 Qy 1230 GATGGGATAA 1239
 Db 1381 GATGGGATAA 1390

RESULT 8

AAF85448
 ID AAF85448 standard; cDNA; 1203 BP.

AC AAF85448;

DT 23-JUL-2001 (first entry)

XX Nucleotide sequence of a rabbit motilin receptor polypeptide.

DE Motilin receptor; gastrointestinal disease; gastric motility disorder;
 KW gastroparesis; irritable bowel syndrome; diarrhoea; 88.

XX Oryctolagus cuniculus.

XX Key Location/Qualifiers
 FH 1..1203
 CDS /tag= a

FT /note= "motilin receptor"

XX WO200132710-A1.

XX 10-MAY-2001.

XX 25-OCT-2000; 2000WO-US29426.

XX 29-OCT-1999; 99US-0162264.

XX (MERI) MERCK & CO INC.

XX Tan C, McKee K;

XX WPI; 2001-343479/36.

DR

P-PSDB; AAB68477.

Novel polypeptides related to dog and rabbit motilin receptor polypeptide, comprising unique regions from dog and motilin receptor amino acid sequence, useful for identifying compounds for treating diarrhoea in humans -

Claim 18; Page 18-19; 42pp; English.

The present sequence encodes a rabbit motilin receptor polypeptide. The specification describes a unique sequence present in exon 1 of the dog motilin receptor, which is not present in human or *Sphaeroides nophilus* 7587 motilin receptor sequences. The unique nucleic acid sequence is useful for measuring the ability of a compound to affect motilin receptor activity. Motilin receptor polynucleotides and polypeptides are used to identify therapeutic compounds which are useful for treating gastrointestinal diseases and disorders such as gastric motility disorders, gastroparesis, irritable bowel syndrome, and diarrhoea.

Sequence 1203 BP; 154 A; 423 C; 403 G; 223 T; 0 other;

Query Match 73.9%; Score 915.6; DB 22; Length 1203;

Best Local Similarity 85.3%; Pred. No. 1.2e-157;

Matches 1058; Conservative 0; Mismatches 144; Indels 36; Gaps 2;

Qy 1 ATGGCGAGCCCTCGAAACGCGACGAGCGCCGCGAGGGGCGGAGCGCGCTGGGCC 60

Db 1 ATGGGACGCCCTTGGAAACGCGACGAGCGCCGCGAGGACGCGGGGAGCGCGCTGGGCC 60

Qy 61 GCCTGCGCGCTTTCGACGAGCGCCGCTGCTCCGCTTTCCCTTGGGGGCGCTGGTGGCG 120

Db 61 GCCTGCGCGCGCTGCGATGAGCGCGCTGCTCCGCTTTCCCTTGGGACGCTGGTGGCT 120

Qy 121 GTACCGCTGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180

Db 121 GTACCGCGCTGTGCTGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180

Qy 181 ATGCTGATCGGGCGCTACCGGGACATGCGGACACCAACCACTTGTACCTGGGCGAGCATG 240

Db 181 CTGCTGATCGGGCGCTACCGGGACATGCGGACACCAACCACTTGTACCTGGGCGAGCATG 240

Qy 241 GCCGTGTCCGACCTACTCATCTGCTGCGGCTGCGCTGCGCTGCGCTGCGCTGCGCGC 300

Db 241 GCCGTGTCCGACCTACTCATCTGCTGCGGCTGCGCTGCGCTGCGCTGCGCTGCGCGC 300

Qy 301 TCGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG 360

Db 301 TCGAGGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 360

Qy 361 TGCACCTAGCCACCGCTGCTGACATGACCGCGCTCAGCGCTCAGCGCTACCTGGCCATC 420

Db 361 TGCACCTAGCCACCGCTGCTGACATGACCGCGCTCAGCGCTCAGCGCTACCTGGCCATC 420

Qy 421 TGCAGCGCGCTCGCGCGCTGCTTGGTCAACCGCGCGCGCTCGCGCGCTCATCGCT 480

Db 421 TGCAGCGCGCTCGCGCGCTGCTTGGTCAACCGCGCGCGCTCGCGCGCTCATCGCT 480

Qy 481 GTGCTCTGGGCGCTGCGGCTGCTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540

Db 481 GCGCTCTGGGCGCTGCGGCTGCTGCTGCGGCGCGCTTCTTCTTCTGCTGCTGCTGCTG 540

Qy 541 CAGGACCGCGCATCTCCGTAGTCCCGGCTCAATGGACACCGCGGATCGCTCTCTCG 600

Db 541 CAGGACCGCGCGTCTTCCGCGCGCGGACCGAACGGTACTGTGCGCTGGACCCCTCG 600

Qy 601 CCTCTGCGCTGCTGCGCGCTCTCTGCTCTCGCGGCGCGCACCGCGCTCCCGCGCTCG 660

Db 601 TCGCGCGCT-----CGGCGCTCCCGCGCTCG 627

Qy 661 GGGCGCGAGACCGCGAGCGCGCGCTGCTTTCAGCGCGGAAATGCGGCGCGAGCGCGCG 720

Db 628 GGGCGCG---GAGCGGAGCGCGGCTCTGTTTCAGCGCGAGTGCAGCGCGAGCGCGCGCG 684

Db 1509 GTGCTGTGGCGCGTCTCTCTGCGCGGTCCCTTTCTTGTCTGTGGTGGCGTCTGAG 1568
 Qy AAA46116
 541 CAGACCCCGGCATCTCCGTAGTCCCGGGCTCAATGGCACCGCGCGATCGCTCTCTCG 600
 Db CAGACCCCGGCATCTCCGTAGTCCCGGGCTCAATGGCACCGCGCGATCGCTCTCTCG 1628
 Qy 601 CTTCTGCTGTGCTGCGCGCTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 660
 Db 1629 CTTCTGCTGTGCTGCGCGCTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 1688
 Qy 661 GGGCCCGAGACCGCGGAGCGCGCGCTGTTAGCGCGGATGCGCGCGAGCGCGCG 720
 Db 1689 GGGCCCGAGACCGCGGAGCGCGCGCTGTTAGCGCGGATGCGCGCGAGCGCGCG 1748
 Qy 721 CAGCTGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 Db 1749 CAGCTGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1808
 Qy 781 CTGTGCTCAGCATCTCTTACGGGCTCATCGGGCGGGAGCTGTGGAGCAGCGCGCGCG 840
 Db 1809 CTGTGCTCAGCATCTCTTACGGGCTCATCGGGCGGGAGCTGTGGAGCAGCGCGCGCG 1868
 Qy 841 CTGCGAGCGCGCGCGCTCGGGCGGGAGAGAGCGCACCGGAGCGCTCGCGCTCTG 900
 Db 1869 CTGCGAGCGCGCGCGCTCGGGCGGGAGAGAGCGCACCGGAGCGCTCGCGCTCTG 1928
 Qy 901 C 901
 Db 1929 C 1929

RESULT 11
 ID AAA46116
 AA AAA46116 standard; cDNA; 2040 BP.
 AC AAA46116;
 XX 22-AUG-2000 (first entry)
 XX Human G protein coupled receptor hGPR38 (V297K) cDNA SEQ ID NO:129.
 XX Human; G protein coupled receptor; GPCR; transmembrane receptor;
 XX identification; agonist; screening; therapeutic; pharmaceutical;
 XX mutant; ss.
 XX Homo sapiens.
 XX Synthetic.
 XX WO200022131-A2.
 XX 20-APR-2000.
 XX 13-OCT-1999; 99WO-US24085.
 XX 13-OCT-1998; 98US-0170496.
 XX 12-NOV-1998; 98US-0108029.
 XX 20-NOV-1998; 98US-0109213.
 XX 27-NOV-1998; 98US-0110060.
 XX 16-FEB-1999; 99US-0120416.
 XX 26-FEB-1999; 99US-0121852.
 XX 12-MAR-1999; 99US-0123944.
 XX 12-MAR-1999; 99US-0123945.
 XX 12-MAR-1999; 99US-0123946.
 XX 12-MAR-1999; 99US-0123948.
 XX 12-MAR-1999; 99US-0123949.
 XX 28-MAY-1999; 99US-0136436.
 XX 28-MAY-1999; 99US-0136437.
 XX 28-MAY-1999; 99US-0136439.
 XX 28-MAY-1999; 99US-0137127.
 XX 28-MAY-1999; 99US-0137131.
 XX 30-JUN-1999; 99US-0137567.
 XX 99US-0141448.

PR 27-AUG-1999; 99US-0151114.
 PR 03-SEP-1999; 99US-0152524.
 PR 29-SEP-1999; 99US-0156633.
 PR 29-SEP-1999; 99US-0156555.
 PR 29-SEP-1999; 99US-0156634.
 XX (AREN-) ARENA PHARM INC.
 XX Behan DP, Lehmann-Bruinema K, Chalmers DT, Chen R, Dang HT;
 PI Gore M, Liaw CW, Lin I, Lowitz K, White C;
 XX WPI; 2000-317986/27.
 DR P-PSDB; AAB02854.
 XX Non-endogenous, human G protein-coupled receptors for screening
 receptor, inverse or partial agonists useful as therapeutic agents -
 Example 2; Page 166-168; 187pp; English.
 XX The present invention describes transmembrane receptors, preferably
 human G protein coupled receptors (GPCR), for which the endogenous
 ligand is unknown (orphan GPCR receptors). More specifically the present
 invention relates to non-endogenous, constitutively activated versions
 of a human GPCR. These non-endogenous human GPCRs can be useful for
 the direct identification of candidate compounds as receptors agonists,
 inverse agonists or partial agonists for use as pharmaceutical agents.
 CC AAA46017 to AAA46126 and AAB02825 to AAB02859 represent sequences used in
 CC the exemplification of the present invention.
 XX Sequence 2040 BP; 350 A; 630 C; 597 G; 463 T; 0 other;
 SQ
 Query Match 72.3%; Score 896.2; DB 21; Length 2040;
 Best Local Similarity 99.7%; Pred. No. 4.2e-154;
 Matches 898; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 ATGGGAGCGCCCTGGAAACGGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
 Db 1 ATGGGAGCGCCCTGGAAACGGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
 Qy 61 GCCTGCGCGCTTGGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
 Db 61 GCCTGCGCGCTTGGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
 Qy 121 GTACACCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
 Db 121 GTACACCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
 Qy 181 ATGCTGATCGGGCGCTACCGGGACATCGGGACATCGGGACATCGGGACATCGGGACATG 240
 Db 181 ATGCTGATCGGGCGCTACCGGGACATCGGGACATCGGGACATCGGGACATCGGGACATG 240
 Qy 241 GCGGTGTCGACCTACTCATCTCTGCTGCGCGCGCTGCTGCGCGCGCTGCTGCGCGCG 300
 Db 241 GCGGTGTCGACCTACTCATCTCTGCTGCGCGCGCTGCTGCGCGCGCTGCTGCGCGCG 300
 Qy 301 TCGCGCGCGCTGGGTGTTTGGGGCGCGCTGCTGCGCGCGCTGCTGCGCGCGCTGCTGCG 360
 Db 301 TCGCGCGCGCTGGGTGTTTGGGGCGCGCTGCTGCGCGCGCTGCTGCGCGCGCTGCTGCG 360
 Qy 361 TGCACCTAGCCACCGCTGTGCAATCGCGCTGTGCAATCGCGCTGTGCAATCGCGCTGTG 420
 Db 361 TGCACCTAGCCACCGCTGTGCAATCGCGCTGTGCAATCGCGCTGTGCAATCGCGCTGTG 420
 Qy 421 TGCAGCGCGCTGCGCGCGCGCTGCTGCTGCGCGCGCTGCTGCGCGCGCTGCTGCGCGCG 480
 Db 421 TGCAGCGCGCTGCGCGCGCGCTGCTGCTGCGCGCGCTGCTGCGCGCGCTGCTGCGCGCG 480
 Qy 481 GTGCTGTGGCGCGTGGCGCTGCTGCTGCGCGCGCTGCTGCGCGCGCTGCTGCGCGCG 540
 Db 481 GTGCTGTGGCGCGTGGCGCTGCTGCTGCGCGCGCTGCTGCGCGCGCTGCTGCGCGCG 540
 Qy 541 CAGGACCCCGGATCTCCGTAGTCCCGGGCTCAATGGCACCGCGCGATCGCTCTCTCG 600
 Db 541 CAGGACCCCGGATCTCCGTAGTCCCGGGCTCAATGGCACCGCGCGATCGCTCTCTCG 600

Db 541 CAGGACCCCGGATCTCGTAGTCCCGGGCTCAATGGCAACGGCGGATCGCTCTCTCG 600
Qy 601 CCTCTCGCTCGTCCCGCTCTCTGCTCTCGGGCGCCACCGCCGCTCCCGCCGTCG 660
Db 601 CCTCTCGCTCGTCCCGCTCTCTGCTCTCGGGCGCCACCGCGTCCCGCCGTCG 660
Qy 661 GGGCCCGAGACCGGAGCGCGCGCTGTTACGCGCGAATGCGGCGCGAGCCCGCG 720
Db 661 GGGCCCGAGACCGGAGCGCGCGCTGTTACGCGCGAATGCGGCGCGAGCCCGCG 720
Qy 721 CAGCTGGGCGCGTCCGCTCATGCTGTTGCTGCTACCGCGCTCTCTCTCTCTCT 780
Db 721 CAGCTGGGCGCGTCCGCTCATGCTGTTGCTGCTACCGCGCTCTCTCTCTCTCT 780
Qy 781 CTGTGCTCAGCATCTCTACGGGCTCATCGGGCGGAGCTGTGGAGCAGCGCGCGCG 840
Db 781 CTGTGCTCAGCATCTCTACGGGCTCATCGGGCGGAGCTGTGGAGCAGCGCGCGCG 840
Qy 841 CTGGAGCGCGCGCTCGGGCGGAGAGAGGCGACCGGCGAGCGTCCGCTCTCTG 900
Db 841 CTGGAGCGCGCGCTCGGGCGGAGAGAGGCGACCGGCGAGCGTCCGCTCTCTG 900
Qy 901 C 901
Db 901 C 901
RESULT 12
AAF85447
ID AAF85447 standard; DNA; 813 BP.
XX AC
XX AAF85447;
XX 23-JUL-2001 (first entry)
DE Nucleotide sequence of a dog motilin receptor exon 1.
XX
KW Motilin receptor; gastrointestinal disease; gastric motility disorder;
KW gastroparesis; irritable bowel syndrome; diarrhoea; ss.
XX Canis sp.
XX
FH Key Location/Qualifiers
FT 1..813
CDS /*tag= a
FT /note= "motilin receptor exon 1"
FT
XX WO200132710-A1.
XX
PD 10-MAY-2001.
XX
PF 25-OCT-2000; 2000WO-US29426.
XX
PR 29-OCT-1999; 99US-0162264.
XX
PA (MERI) MERCK & CO INC.
XX
XX Tan C, McKee K;
XX WPI; 2001-343479/36.
DR P-PSDB; AAB68476.
XX
XX Novel polypeptides related to dog and rabbit motilin receptor
PT polypeptide, comprising unique regions from dog and motilin receptor
PT amino acid sequence, useful for identifying compounds for treating
PT diarrhoea in humans
XX
PS Claim 14; Page 17-18; 42pp; English.
XX
CC The present sequence represents exon 1 of a dog motilin receptor gene.
CC The specification describes an unique sequence present in exon 1 of
CC the motilin receptor, which is not present in human or Sphaeroides
CC nephelus 75E7 motilin receptor sequences. The unique nucleic acid

CC sequence is useful for measuring the ability of a compound to affect
CC motilin receptor activity. Motilin receptor polynucleotides and
CC polypeptides are used to identify therapeutic compounds which are
CC useful for treating gastrointestinal diseases and disorders such as
CC gastric motility disorders, gastroparesis, irritable bowel syndrome,
CC and diarrhoea.

XX SQ Sequence 813 BP; 65 A; 316 C; 305 G; 127 T; 0 other;

Query Match 44.1%; Score 546; DB 22; Length 813;

Best Local Similarity 80.5%; Pred. No. 1.9e-90;

Matches 705; Conservative 0; Mismatches 105; Indels 66; Gaps 3;

Qy 25 GACGCCCCAGGGGCGCGGAGCCCGCTGGCGCCGCTGGCGCTTGGAGAGAGCGC 84

Db 4 GGGCGCCCGGGAACAGACGACGGCGCGAGGGCGCAGCTGCCGTGCGAGAGCGC 63

Qy 85 CGCTGCTCGCCCTTCCCTTGGGGGCGCTGGTCCGGTGACCGCTGTGCTGTGCTG 144

Db 64 CTGTGCTCGCCCTTCCCTTGGGGGCGCTGGTCCGGTGACCGCTGTGCTGTGCTG 123

Qy 145 TTGCTGCTCGGGGTGAGCGCAACGTGTACCGTGTATGTATCGGGCGCTACCGGAC 204

Db 124 TTGCTGCTCGGGGTGAGCGCAACGTGTATCGGGCGCTACCGGCGCTACCGGAC 183

Qy 205 ATCGGACACCAACCACTTGTATCTGGGAGCATGCGCGCTGTCGACCTACTCATCT 264

Db 184 ATCGGACACCAACCACTTGTATCTGGGAGCATGCGCGCTGTCGACCTACTCATCT 243

Qy 265 CTGGGCTGCGCTTGGAGCTGTACCGCTCTGGGGCTCGGGCGCTGGGGCTGTCGG 324

Db 244 CTGGGCTGCGCTTGGAGCTGTACCGCTCTGGGGCTCGGGCGCTGGGGCTGTCGG 303

Qy 325 CTGCTCTGCGCGCTGCTCCCTCTAGTGGGAGGGCTGACCTACGCGACGCTGTCAC 384

Db 304 CTGCTCTGCGCGCTGCTCCCTCTAGTGGGAGGGCTGACCTACGCGACGCTGTCAC 363

Qy 385 ATGACCGCGCTCAGCGCTGAGCGCTACTTGGCGCATCTGCGCGCCGCTCCGCGCG 444

Db 364 GTGACGGCGCTGAGCGCTGAGCGCTACTTGGCGCATCTGCGCGCCGCTCCGCGCG 423

Qy 445 TTGCTCACCGCGCGCGCTGCGCGCTCATGCTGCTGTGCTGCGGGCTGGCGCTGCTC 504

Db 424 CTGCTGTCGCGCGCGCGCGCTCATGCTGCGCGCTCTGCGGGCTGGCGCTGCTG 483

Qy 505 TCTGCGGTCCTCTTCTTCTGCTGGTGGGCTGAGCAGGACCCCGGCTCTCGGTAGTC 564

Db 484 TCGGCGCGCGCTTCTTCTTCTGCTGGTGGGCTGAGCAGGACCGGGGCTGCTG 534

Qy 565 CCGGCGCTCAATGGCACCGCGCGGATCGCTCTCTGCGCTCTGCGCGCTCTCTCTC 624

Db 535 CCGGCGCTCAATGGCACCGCGCGGATCGCTCTCTGCGCTCTGCGCGCTCTCTCTC 557

Qy 625 TGGCTCTGCGGGCGCGCACCGCGTCCCGCTGCGGGCGCGAGACCGGAGAGCGCGCG 684

Db 558 -----GCTGGCGCGGGCGCGCTCCCGCGCGGGGCGCGAG-----GCG 597

Qy 685 GCGCTGTTCAGCGCGGAATGCGGGCGAGCGCGCGCGAGCTGGGCGGCTGCTGTCATG 744

Db 598 GCGCTGTTCAGCGCGGAATGCGGGCGAGTCCCGCGCGAGCTGGGCGGCTGCTGTCATG 657

Qy 745 CTGTGGGTCAACCGCGCTACTTCTTCTGCGCTTCTGCTGCTGCTGCTGCTGCTG 804

Db 658 CTGTGGGTCAACCGCGCTACTTCTTCTGCGCTTCTGCTGCTGCTGCTGCTGCTG 717

Qy 805 CTCAATCGGGCGGAGCTGTGGAGCAGCGCGCGCTGCGAGGCGCGCGCTCGGGG 864

Db 718 CGCATCGGCGGAGCTGTGGAGGCGCGCGCGCTGCGGGCGCGCGCTCGGGGCTCGGG 777

Qy 865 CGGAGAGAGCGCCACCGCGAGACCGTCCGCTGCTG 900

Db 778 CGCAGCGGGGCCACCGCGAGCGCGCTCCGCTGCTG 813

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 31, 2003, 21:43:42 ; Search time 87.6584 Seconds
(without alignments)
6238.690 Million cell updates/sec

Title: US-09-719-485-2

Perfect score: 1239

Sequence: 1 atgggcagccctggaacgg.....acgtgaagacgatgggataa 1239

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PTUS COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	283	22.8	283	3	US-08-993-088A-4
2	283	22.8	283	4	US-08-993-424B-4
3	283	22.8	283	4	US-09-603-680-4
4	234.8	19.0	1063	3	US-09-077-675A-1
5	234.8	19.0	1063	4	US-09-077-674-1
6	230	18.6	1029	3	US-09-077-675A-4
7	230	18.6	1029	4	US-09-077-674-4
8	229.8	18.5	1092	3	US-09-077-675A-15
9	229.8	18.5	1092	4	US-09-077-674-15
10	229	18.5	250	4	US-09-016-434-359
11	228.8	18.5	1122	3	US-09-077-675A-9
12	228.8	18.5	1122	4	US-09-077-674-9
13	228.4	18.4	3129	3	US-08-077-675A-14
14	228.4	18.4	3129	4	US-09-077-674-14
15	227.8	18.4	1088	3	US-09-077-675A-6
16	227.8	18.4	1088	4	US-09-077-674-6
17	227.8	18.4	1101	4	US-09-016-434-1148
18	227.8	18.4	1101	4	US-09-170-496D-87
19	227.8	18.4	1101	4	US-09-170-496D-209
20	153	12.3	836	3	US-09-077-675A-11
21	153	12.3	836	4	US-09-077-674-11
22	134.8	10.9	1248	4	US-09-545-944-1
23	132	10.7	1575	3	US-08-858-876A-1
24	132	10.7	1575	3	US-09-472-880-1
25	125.4	10.1	1342	3	US-08-832-399-1
26	125.4	10.1	1342	3	US-09-372-498-1
27	123	9.9	4080	4	US-09-016-434-1346

Sequence 3, Appli
Sequence 3, Appli
Sequence 1366, Ap
Sequence 1, Appli
Sequence 12, Appli
Sequence 113, App
Sequence 223, App
Sequence 3, Appli
Sequence 6, Appli
Sequence 6, Appli
Sequence 27, Appli
Sequence 3, Appli
Sequence 18, Appli
Sequence 19, Appli
Sequence 18, Appli
Sequence 18, Appli
Sequence 19, Appli

28 122.8 9.9 1529 3 US-08-958-876A-3
29 122.8 9.9 1529 3 US-09-472-880-3
30 122 9.8 1285 4 US-09-016-434-1366
31 120.4 9.7 1233 4 US-09-200-090-1
32 120 9.7 1535 4 US-09-568-680-12
33 119 9.6 1212 4 US-09-170-496D-113
34 119 9.6 1212 4 US-09-170-496D-223
35 115.4 9.3 1176 4 US-09-200-090-3
36 109.8 8.9 1164 3 US-08-993-088A-6
37 109.8 8.9 1164 4 US-08-993-424B-6
38 109.8 8.9 1164 4 US-09-603-680-6
39 109.8 8.9 1365 4 US-08-899-112B-27
40 108.2 8.7 1219 4 US-08-981-700A-3
41 106.8 8.6 1116 3 US-08-993-088A-18
42 106.8 8.6 1116 4 US-08-993-088A-19
43 106.8 8.6 1116 4 US-08-993-424B-18
44 106.8 8.6 1116 4 US-09-603-680-18
45 106.8 8.6 1116 4 US-09-603-680-19

ALIGNMENTS

RESULT 1
US-08-993-088A-4
; Sequence 4, Application US/08993088A
; Patent No. 6287855
; GENERAL INFORMATION:
; APPLICANT: Tan, Carina
; APPLICANT: Sullivan, Kathleen
; TITLE OF INVENTION: GALANIN RECEPTOR GALR2 AND
; TITLE OF INVENTION: NUCLEOTIDES ENCODING SAME
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FASTSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/993,088A
; FILING DATE: 18-DEC-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/033,851
; FILING DATE: 27-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Heber, Sheldon O.
; REGISTRATION NUMBER: 38,179
; REFERENCE/DOCKET NUMBER: 19846
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1958
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 283 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Other
; FEATURE:
; NAME/KEY: Other
; LOCATION: 1...283
; OTHER INFORMATION: cdna probe
; US-08-993-088A-4

Query Match 22.8%; Score 283; DB 3; Length 283;
Best Local Similarity 100.0%; Pred. No. 4.8e-48;
Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 206 TGGGACACCAACTTGTACCTGGGAGCAGTGGCCGTGTCGACCTACTATCTGTC 265
DB 1 TGGGACACCAACTTGTACCTGGGAGCAGTGGCCGTGTCGACCTACTATCTGTC 60

QY 266 TGGGCTGCGCTTGCACCTGTACCGCTCTGGGCTCGCGCCCTGGGTGTTGGGGCGC 325
DB 61 TGGGCTGCGCTTGCACCTGTACCGCTCTGGGCTCGCGCCCTGGGTGTTGGGGCGC 120

QY 326 TGCTCTCGCCCTGTCCCTCTACGTGGGCGAGGCTGCACCTAGCCACGCTGTCGACA 385
DB 121 TGCTCTCGCCCTGTCCCTCTACGTGGGCGAGGCTGCACCTAGCCACGCTGTCGACA 180

QY 386 TGACCGGCTCAGGCTCGAGGCTACCTGGCCATCTGCGCGCGCTCGCGCGCGGTCT 445
DB 181 TGACCGGCTCAGGCTCGAGGCTACCTGGCCATCTGCGCGCGCTCGCGCGCGGTCT 240

QY 446 TGGTACCCGCGCGCGCTCGCGCGCTCATCGCTGTGCTCTG 488
DB 241 TGGTACCCGCGCGCGCTCGCGCGCTCATCGCTGTGCTCTG 283

RESULT 2

US-08-993-424B-4
; Sequence 4, Application US/08993424B
; Patent No. 6337206
; GENERAL INFORMATION:
; APPLICANT: Tan, Carina
; APPLICANT: Kolakowski, Lee P., Jr.
; TITLE OF INVENTION: MOUSE GALANIN RECEPTOR GALR2 AND
; TITLE OF INVENTION: NUCLEOTIDES ENCODING SAME
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/993,424B
; FILING DATE: 18-DEC-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/033,851
; FILING DATE: 27-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Heber, Sheldon O.
; REGISTRATION NUMBER: 38,179
; REFERENCE/DOCKET NUMBER: 19846NP2
; TELEPHONE: 732-594-1958
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 283 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Other
; FEATURE:
; NAME/KEY: Other
; LOCATION: 1...283
; OTHER INFORMATION: cdna probe

US-08-993-424B-4

Query Match 22.8%; Score 283; DB 4; Length 283;
Best Local Similarity 100.0%; Pred. No. 4.8e-48;
Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 206 TGGGACACCAACTTGTACCTGGGAGCAGTGGCCGTGTCGACCTACTATCTGTC 265
DB 1 TGGGACACCAACTTGTACCTGGGAGCAGTGGCCGTGTCGACCTACTATCTGTC 60

QY 266 TGGGCTGCGCTTGCACCTGTACCGCTCTGGGCTCGCGCCCTGGGTGTTGGGGCGC 325
DB 61 TGGGCTGCGCTTGCACCTGTACCGCTCTGGGCTCGCGCCCTGGGTGTTGGGGCGC 120

QY 326 TGCTCTCGCCCTGTCCCTCTACGTGGGCGAGGCTGCACCTAGCCACGCTGTCGACA 385
DB 121 TGCTCTCGCCCTGTCCCTCTACGTGGGCGAGGCTGCACCTAGCCACGCTGTCGACA 180

QY 386 TGACCGGCTCAGGCTCGAGGCTACCTGGCCATCTGCGCGCGCTCGCGCGCGGTCT 445
DB 181 TGACCGGCTCAGGCTCGAGGCTACCTGGCCATCTGCGCGCGCTCGCGCGCGGTCT 240

QY 446 TGGTACCCGCGCGCGCTCGCGCGCTCATCGCTGTGCTCTG 488
DB 241 TGGTACCCGCGCGCGCTCGCGCGCTCATCGCTGTGCTCTG 283

RESULT 3

US-09-603-680-4
; Sequence 4, Application US/09603680
; Patent No. 6544753
; GENERAL INFORMATION:
; APPLICANT: Tan, Carina
; APPLICANT: Sullivan, Kathleen
; TITLE OF INVENTION: GALANIN RECEPTOR GALR2 AND
; TITLE OF INVENTION: NUCLEOTIDES ENCODING SAME
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/603,680
; FILING DATE: 26-Jun-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/033,851
; FILING DATE: 27-DEC-1996
; APPLICATION NUMBER: 08/993,088
; FILING DATE: 18-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Heber, Sheldon O.
; REGISTRATION NUMBER: 38,179
; REFERENCE/DOCKET NUMBER: 19846 CA
; TELEPHONE: 732-594-1958
; TELEFAX: 732-594-4720
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 283 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Other

```
;
; FEATURE:
; NAME/KEY: Other
; LOCATION: 1...283
; OTHER INFORMATION: cdna probe
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-603-680-4

Query Match      22.8%; Score 283; DB 4; Length 283;
Best Local Similarity 100.0%; Pred. No. 4.8e-48;
Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 206 TGGGACACACCAACTTGTACCTGGGAGCATGGCGGTTCGACCTACTCATCTGC 265
Db 1 TGGGACACACCAACTTGTACCTGGGAGCATGGCGGTTCGACCTACTCATCTGC 60

Qy 266 TCGGCTCGCGCTTCGACCTGTACCGCTCTTGGCGCTCGCGCGCTGGGTTCGGGCGC 325
Db 61 TCGGCTCGCGCTTCGACCTGTACCGCTCTTGGCGCTCGCGCGCTGGGTTCGGGCGC 120

Qy 326 TGCTCTCGCGCTTCCTCTTACCTGGGAGGCTGCACCTACGCCACGCTGCTGCACA 385
Db 121 TGCTCTCGCGCTTCCTCTTACCTGGGAGGCTGCACCTACGCCACGCTGCTGCACA 180

Qy 386 TGACCGGCTCAGGCTCGAGGCTTACCTGGCCATCTGCGCGCGCTCGCGCGCTCT 445
Db 181 TGACCGGCTCAGGCTCGAGGCTTACCTGGCCATCTGCGCGCGCTCGCGCGCTCT 240

Qy 446 TGGTACCGCGCGCGCTCGCGCGCTGCTGCTGCTGCTG 488
Db 241 TGGTACCGCGCGCGCTCGCGCGCTGCTGCTGCTGCTG 283

RESULT 4
US-09-719-485-2
; Sequence 1, Application US/09077675A
; Patent No. 624129
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1063 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-09-077-675A-1

Query Match      19.0%; Score 234.8; DB 3; Length 1063;
Best Local Similarity 69.3%; Pred. No. 2.3e-36;
Matches 320; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

Qy 92 CGCCCTTTCCCTTGGGGCGCTGGTGGCGGTGACCGCTGATGCTGATCGGGCGCTACCGGACATGGGGA 151
Db 69 CGCTTTTCCCTTGGGGCGCTGGTGGCGGTGACCGCTGATGCTGATCGGGCGCTACCGGACATGGGGA 128

Qy 152 TCGGGGTGAGCGGCAACGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 211
Db 129 TGGGTATCGGGGCAACGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 188

Qy 212 CCACCAACAACTTGTACTGGGAGCATGGCGGTGCTCGGACCTACTCATCTGCTCGGGC 271
Db 189 CCACCAACAACTTGTACTGGGAGCATGGCGGTGCTCGGACCTACTCATCTGCTCGGGC 248

Qy 272 TCGGCTTGCACCTTACCGCTCTGGCGCTCGCGCGCTCGCGCGCTCGCGCGCTCGCGCGCTCTCT 331
Db 249 TCGGCTTGCACCTTACCGCTCTGGCGCTCGCGCGCTCGCGCGCTCGCGCGCTCGCGCGCTCTCT 308

Qy 332 GCGGCTTGCACCTTACCGCTCTGGCGCTCGCGCGCTCGCGCGCTCGCGCGCTCGCGCGCTCTCT 391
Db 309 GCAAACTCTTCCAGTTCGTTAGGAGAGCTGCACTAGCCACAGTCTTACCATCACCG 368

Qy 392 CGCTCAGCGTCGAGCGCTTACCTGCGCATCTGCGCGCTCGCGCGCTCGCGCGCTCTTGGTCA 451
Db 369 CGCTGAGCGTCGAGCGCTTACCTGCGCATCTGCGCGCTCGCGCGCTCGCGCGCTCTTGGTCA 428

Qy 452 CCGGCGCGCGCGCTCGCGCGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 511
Db 429 CCAAGGCGCGGTAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 488

Qy 512 GTCCCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 553
Db 489 GGCCCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 530

RESULT 5
US-09-077-674-1
; Sequence 1, Application US/09077674
; Patent No. 653114
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR FAMILY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,674
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1063 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-09-077-675A-1
```

CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19589P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1063 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-077-674-1

Query Match 19.08; Score 234.8; DB 4; Length 1063;
Best Local Similarity 69.3; Pred. No. 2.3e-38;
Matches 320; Conservative 0; Mismatches 142; Indels 0; Gaps 0;
QY 92 CGCCCTTTCCCTGGGGGCGTGTGCGCGGTGACCGCTGTGTGCTGTGCTGTGCTGTGCTG 151
DB 69 CGCTCTTCCCAACGCCGCTGTGGGGGGGTGACCGGCACCTGGTGGCGCTCTTCTGTG 128
QY 152 TCGGGGTGAGCGGCAACGCTGTGACCGGTGTGATGCTGTGATCGGGCGCTACCGGACATGCGGA 211
DB 129 TGGGTATCGCGGGCAACCTGCTACGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 188
QY 212 CCACCACCACTTGTACTTGGGCGAGATGGCGGTGTGCGGCTGTGCTGTGCTGTGCTGTG 271
DB 189 CCACCACCACTTGTACTTGGGCGAGATGGCGGTGTGCGGCTGTGCTGTGCTGTGCTGTG 248
QY 272 TGGGTATCGCGGGCAACCTGCTACGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 331
DB 249 TGGGTATCGCGGGCAACCTGCTACGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 308
QY 332 GCGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 391
DB 309 GCAAACTCTTCCAGTTCGTTAGCGAGAGCTGCACCTAGCCACACAGTCTCACCATCACCG 368
QY 392 CGCTCAGCGTGCAGCGCTACTTGGCCATCTGCGCGCCGCTTCCGCGCCGCGCTTCTGTG 451
DB 369 CGCTCAGCGTGCAGCGCTACTTGGCCATCTGCTTCCGCGCTTCCGCGCCGCGCTTCTGTG 428
QY 452 CCGGCGCGCGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 511
DB 429 CCAAGGCGCGGTTAAAGCTGT 488
QY 512 GTCCCTTTCTTTGTTCTTCTGTTGGGGGTGAGCAGGACCCCGGCA 553
DB 489 GGCCCATCTTTCGTGCTGTGCTGCGAGTGGAGCATGATACGGCA 530

RESULT 6
US-09-077-675A-4
Sequence 4, Application US/09077675A
Patent No. 6242199
GENERAL INFORMATION:
APPLICANT: Fai, Lee-Yuh
APPLICANT: Feighner, Scott C.
APPLICANT: Howard, Andrew D.
APPLICANT: Pong, Sheng-Shung
APPLICANT: Van Der Ploeg, Leonardus H.T.
TITLE OF INVENTION: RECEPTOR ASSAY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.

CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/077,675A
FILING DATE: 3-JUN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19590P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1029 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-077-675A-4

Query Match 18.68; Score 230; DB 3; Length 1029;
Best Local Similarity 68.6; Pred. No. 2e-37;
Matches 317; Conservative 0; Mismatches 145; Indels 0; Gaps 0;
QY 92 CGCCCTTTCCCTGGGGGCGTGTGCGCGGTGACCGCTGTGTGCTGTGCTGTGCTGTGCTG 151
DB 266 CGCTCTTCCCAACGCCGCTGTGGGGGGGTGACCGGCACCTGGTGGCGCTCTTCTGTG 325
QY 152 TCGGGGTGAGCGGCAACGCTGTGACCGGTGTGATGCTGTGATCGGGCGCTACCGGACATGCGGA 211
DB 326 TGGGTATCGCGGGCAACCTGCTCAGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 385
QY 212 CCACCACCACTTGTACTTGGGCGAGATGGCGGTGTGCGGCTGTGCTGTGCTGTGCTGTG 271
DB 386 CCACCACCACTTGTACTTGGGCGAGATGGCGCTTCTCCGAACCTACTCATCTTCTCTGCA 445
QY 272 TGGGTATCGCGGGCAACCTGCTCAGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 331
DB 446 TGGGTATCGCGGGCAACCTGCTCAGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 505
QY 332 GCGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 391
DB 506 GCAAACTCTTCCAGTTCGTTAGCGAGAGCTGCACCTAGCCACACAGTCTCACCATCACCG 565
QY 392 CGCTCAGCGTGCAGCGCTACTTGGCCATCTGCGCGCCGCTTCCGCGCCGCGCTTCTGTG 451
DB 566 CGCTCAGCGTGCAGCGCTACTTGGCCATCTGCTTCCGCGCTTCCGCGCCGCGCTTCTGTG 625
QY 452 CCGGCGCGCGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 511
DB 626 CCAAGGCGCGGTTAAAGCTGT 685
QY 512 GTCCCTTTCTTTGTTCTTCTGTTGGGGGTGAGCAGGACCCCGGCA 553
DB 686 GGCCCATCTTTCGTGCTGTGCTGCGAGTGGAGCATGATACGGCA 727

RESULT 7
US-09-077-674-4
Sequence 4, Application US/09077674

224	DB	CA	CC	ACCA	CA	ACCT	T	CA	CT	GC	AG	AT	GG	CC	T	T	CG	AT	CT	GC	CT	CA	T	T	C	T	CT	GC	CA	283		
272	QY	TG	CG	TT	GC	AT	TA	AC	CG	CT	CT	GG	CG	CT	CG	CG	CC	CT	GG	TT	TC	GG	GC	CG	CT	GC	T	CT	CT	331		
284	DB	TG	CG	CT	TG	GA	CC	T	CG	CT	CT	TG	GA	T	AC	CG	CC	CT	GA	AT	TC	GG	GC	AC	CT	GC	T	CT	343			
332	QY	GC	CG	CT	GC	CT	T	AG	CT	GG	GC	AG	GG	CT	GC	AC	CT	AG	GC	CA	CG	CT	GC	CA	AT	GA	CC	GC	391			
344	DB	GC	AA	CT	T	T	CC	AG	T	T	T	T	GC	AG	GA	GT	GC	AC	T	AG	CG	CA	CG	CT	C	T	CA	CC	AT	CA	CC	403
392	QY	CG	CT	AG	CT	GC	AG	CG	CT	AC	CT	GC	GC	CA	T	GC	CG	CC	CG	CT	CG	CG	CG	CG	CT	TC	TC	GC	TC	451		
404	DB	CG	CT	AG	CT	GC	AG	CG	CT	AC	CT	GC	GC	CA	T	GC	CG	CC	CG	CT	CG	CG	CG	CG	CT	TC	TC	GC	TC	463		
452	QY	CC	CG	CG	CG	CG	CT	CG	CG	CG	CT	CA	CT	GC	T	GC	CT	GC	GG	CG	CT	GC	CT	CT	CT	CT	CT	CT	CT	511		
464	DB	CT	AA	GG	CG	CG	CT	GA	CT	GG	T	CA	T	CT	T	GC	AT	CT	GG	CG	CT	GG	CT	TC	TC	GC	AG	CG	CG	523		
512	QY	GT	CC	CT	TT	TC	CT	GG	TC	GG	CG	TC	AG	CA	GA	CC	CC	CG	CA	TC	TC	CG	T	AG	T	CC	CG	GC	571			
524	DB	GG	CC	CA	T	CT	CG	T	GC	T	GG	CG	TC	AG	CA	GA	AA	CG	CA	-----	-----	-----	-----	-----	-----	-----	-----	-----	565			
572	QY	TC	AA	TG	GC	AC	CG	CG	GA	TC	GC	CT	CT	CG	CT	CT	CG	CC	CT	CT	CG	CC	CT	CT	CT	CT	CT	CT	631			
566	DB	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	572				
632	QY	CG	CG	CG	CG	CA	CG	CG	CT	CC	CG	CG	TC	GG	GC	CG	AG	CA	CG	CG	AG	CG	CG	CG	CG	CT	GC	T	691			
573	DB	CG	GG	GA	CA	CA	GA	AT	CG	CG	CC	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	625				
692	QY	TC	AG	CG	CA	AT	CG	CG	CG	CG	CG	CG	CA	GT	GG	CG	CG	TC	CG	T	GC	T	CA	TC	T	GC	T	751				
626	DB	TC	AC	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	643				
752	QY	TC	AC	CA	CG	CG	CT	ACT	TT	CT	CG	CC	TT	CT	GC	CT	C	AG	AT	CT	CT	AG	CG	CT	CA	TC	GC	811				
644	DB	TG	TC	CA	AG	CG	CT	TT	CT	TT	CT	TA	CC	GG	CT	CT	TC	GC	CT	CA	CT	GT	GT	CT	TA	CA	TC	GC	703			
812	QY	GG	CG	GA	CT	TC	GG	AG	CA	CG	CG	CG	CG	CT	GC	AG	CC	CG	CG	CG	CT	CG	GG	CG	CG	CG	CG	871				
704	DB	GG	AG	GA	CT	AT	GG	CG	AG	CA	CG	CG	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	760					
872	QY	GA	GC	CA	CG	CG	CA	CG	CT	CC	CG	GT	CC	TC	GT	GG	T	CT														

RESULT 9

US-09-077-674-15

; Sequence 15, Application US/09077674

; Patent No. 6531314

GENERAL INFORMATION:

APPLICANT: Arena, Joseph P.

APPLICANT: Cully, Doris F.

APPLICANT: Feigner, Scott D.

APPLICANT: Howard, Andrew D.

; APPLICANT: Liberator, Paul A.

[illegible]

```

RESULT 10
US-09-016-434-359
; Sequence 359, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Sellhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:

```

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 359:
SEQUENCE CHARACTERISTICS:
LENGTH: 250 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: TFP1NOT01
CLONE: 2018536
US-09-016-434-359

Query Match	18.5%	Score 229;	DB 4;	Length 250;
Best Local Similarity	98.8%;	Pred. No. 2.6e-37;		
Matches 240;	Conservative 0;	Mismatches 2;	Indels 1;	Gaps 1;
Qy	997	TCTCAGTACTTTAAACATCGTGCCTCTGCAACTTTTCTATCTGAGCGCATCTATCAACCCA	1056	
Db	1	TCTCAGTACTTTAAACATCGTGCCTCTGCAACTTTTCTATCTGAGCGCNTCTATCAACCCA	60	
Qy	1057	ATCCTCTACAACTCTATTCCAAGAGAGTACAGAGCGCGCGCTTTAAACTGCTGCTCGCA	1116	
Db	61	ATCCTCTACAACTCTATTCCAAGAGAGTACAGAGCGCGCGCTTTAAACTGCTGCTCGCA	120	
Qy	1117	AGGAAGTCCAGGCGGAGAGGCTTCCACAGAAAGCAGGAGACACTCGGGGGGAAGTTGCAGGG	1176	
Db	121	AGNAGTCCAGGCGGAGAGG- TTCCACAGAAAGCAGGAGACACTCGGGGGGAAGTTGCAGGG	179	
Qy	1177	GACACTGGAGGAGACACGGTGGGCTACCCGAGACAAGCGCTAAACGTGAAGACGATGGGA	1236	
Db	180	GACACTGGAGGAGACACGGTGGGCTACCCGAGACAAGCGCTAACGTGAAGACGATGGGA	239	
Qy	1237	TAA	1239	
Db	240	TAA	242	

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```


QY 333 CCGCTCTCCCTCTACCTGGGCGAGGGCTGACCTACGCGACCGTGTGTGCACATGACCGC 392
Db 600 CAAACTCTTCAATTCGTAGTGAGAGCTGCACCTAGCGACGGTCTCACCATCACAGC 659
QY 393 GCTCAGGCTGAGCGCTACCTGGCCATCTCGCGCCGCTCGCGCCGCTCTTGGTTCAC 452
Db 660 GCTGAGGCTGAGCGCTACCTGGCCATCTCGCGCCGCTCGCGCCGCTCTTGGTTCAC 719
QY 453 CCGCGCGCGGCTCGCGCGCTCATCGCTGTGCTCTGGGCGGCTGGCGCTGTCTCTGCGG 512
Db 720 CAAGGGCGGGTGAAGCTGGTTCATCTTCTGCTCATCTGGCGCGTGGCTTCTGCGAGCGCG 779
QY 513 TCCCTTTCTTGTCTGTGGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 571
Db 780 GCGCATCTTCTGTAGTGTGGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 839
QY 572 TCAATGACAGCGCGGATCGCTCTCGCTCTCGCTCTCGCTCTCGCTCTCGCTCTCGCTCT 631
Db 840 CGAGTGGCGGCGGAG 899
QY 632 CGCGGGCGGCGGAG 691
Db 900 GTCCAGCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 959
QY 692 TCAGCGCGGAG 751
Db 960 CAGGAAGCTGTGGCGGAG 1019
QY 752 TCAGCGCGGAG 811
Db 1020 GAACCAAG 1079
QY 812 GCGCGGAGCTGT 823
Db 1080 CGCGGGTCTTAT 1091

RESULT 13

US-09-077-675A-14
; Sequence 14, Application US/09077675A
; Patent No. 6242199
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3129 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-077-675A-14

Query Match 18.4%; Score 228.4; DB 3; Length 3129;
Best Local Similarity 68.4%; Pred. No. 5e-37;
Matches 316; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

QY 92 CGCCCTTTCCCTCGGGGCGCTGTCGGGCTGACCGGTGATGCTGATCGGGCGCTACCGGACATGCGGA 151
Db 104 CGCTGTTCCCGCTCGCTGCGGCTGACCGGCTGACCGGCTGCGGCTGCGGCTGCGGCTGCGG 163
QY 152 TCGGGGTGAGCGCAACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 211
Db 164 TGGGCATCTCAGGCAACCTGCTCCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 223
QY 212 CCACCAACAACTTGTACCTGGGCGAGATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 271
Db 224 CCACCAACAACTTGTACCTGGGCGAGATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 283
QY 272 TGGGCTGAGACCTGTACCGGCTGTCGGGCTGCGGGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCT 331
Db 284 TGGGCTGAGACCTGTGTCGGCTGTCGGGCTGTCGGGCTGTCGGGCTGTCGGGCTGTCGGGCTGCT 343
QY 332 GCGGCTGTCCCTCTACGTGGGCGAGGGCTGACCTACGCGACGCTGCTGCTGCTGCTGCTGCTGCT 391
Db 344 GCAAACTTTCAGGTCGAGGAGCTGACCTACGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 403
QY 392 CGCTCAGCGTCGAGCGCTACCTGCGCATCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCTGCT 451
Db 404 CGCTGAGCGTCGAGCGCTACTTGGCCATCTGCTTCCCTCTGCGGGCCAAAGTGGTGGTCA 463
QY 452 CCGGGCGCGGCTCGCGGCTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 511
Db 464 CTAAGGCGCGCTGAAAGCTGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 523
QY 512 GTCCCTTCTTGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 553
Db 524 GGCCCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 565

RESULT 14

US-09-077-674-14
; Sequence 14, Application US/09077674
; Patent No. 6531314
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris P.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR FAMILY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/077,674
 FILING DATE: 3-JUN-1998
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Cocuzzo, Anna L.
 REGISTRATION NUMBER: 42,452
 REFERENCE/DOCKET NUMBER: 19589P
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 732-594-1273
 TELEFAX: 732-594-4720
 TELEX:
 INFORMATION FOR SEQ ID NO: 14:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3129 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cdna
 US-09-077-674-14

Query Match 18.4%; Score 228.4; DB 4; Length 3129;
 Best Local Similarity 68.4%; Pred. No. 5e-37;
 Matches 316; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

Qy	92	CGCCCTTTCCCTGGGGCGCTGGTGGCGGTGACCGCTGTGCTGTGCTGTGCTGTGCTG	151
Db	104	CGCTGTTCCCGCTCCGCTGCTGGCAGCGCTACCGCCACCTGGTGGCGCTCTTCGTGG	163
Qy	152	TCGGGGTGGGGCAACGTGGTGAACCGTGAATGCTGATCGGGCGCTACCGGGAATGCGGA	211
Db	164	TGGGCATCTCAGGCAACCTGCTCACTATGCTGTGTGTGCTCCCGCTTCCGGAGCTGGCA	223
Qy	212	CCACCACCACTTGACTGGGAGCATGCGGTGTCGGACCTACTCATCTGCTCGGGC	271
Db	224	CCACCACCACTTGACTGGGAGCATGCGGTGTCGGACCTACTCATCTGCTCGGGC	283
Qy	272	TGCGGTTGCACTGTACCGCTCTGGCGCTCGCGGCTCGGGTGTGCGGCGCGTGTCT	331
Db	284	TGCGGTTGCACTGTACCGCTCTGGCGCTCGCGGCTCGGGTGTGCGGCGCGTGTCT	343
Qy	332	GCAGCTGTCTTACGTGGGCGAGGGTGCACCTACGCGACGCTGCTGCAATGACCG	391
Db	344	GCAAACTCTTCCAGTTTGTGAGCGAGAGCTGCACCTACGCGACGCTCTCACCATCACCG	403
Qy	392	CGCTCAGGTCGAGCGCTACTGCGCATCTGCGCGCGCTCGCGCGCGCTCTTGTGCA	451
Db	404	CGCTCAGGTCGAGCGCTACTGCGCATCTGCGCGCGCTCGCGCGCGCTCTTGTGCA	463
Qy	452	CCCGGCGCGGCTCGCGCGCTCATCGCTGTGCTGTGCGCGCTGCGCGCTCTCTCGCG	511
Db	464	CTAAGGCGCGGCTGAAGCTGGTCACTCTGTCATCTGCGCGCTGCGCGCTCTCTCGCG	523
Qy	512	GTCCCTTTCTTGTCTCTGTTGGGGTGCAGCAGGACCCCGGCA	553
Db	524	GGCCCATCTTCTGCTGTGGGGTGGAGCAGCAAAACCGCA	565

RESULT 15
 US-09-077-675A-6
 Sequence 6, Application US/09077675A
 Patent No. 6242199
 GENERAL INFORMATION:
 APPLICANT: Pai, Lee-Yuh
 APPLICANT: Feigner, Scott C.
 APPLICANT: Howard, Andrew D.
 APPLICANT: Pong, Sheng-Shung
 APPLICANT: Van Der Ploeg, Leonardus H.T.

TITLE OF INVENTION: RECEPTOR ASSAY
 NUMBER OF SEQUENCES: 16
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Merck & Co., Inc.
 STREET: P.O. Box 2000, 126 E. Lincoln Ave.
 CITY: Rahway
 STATE: NJ
 COUNTRY: USA
 ZIP: 07065-0900
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/077,675A
 FILING DATE: 3-JUN-1998
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Cocuzzo, Anna L.
 REGISTRATION NUMBER: 42,452
 REFERENCE/DOCKET NUMBER: 19590P
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 732-594-1273
 TELEFAX: 732-594-4720
 TELEX:
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1088 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cdna
 US-09-077-675A-6

Query Match 18.4%; Score 227.8; DB 3; Length 1088;
 Best Local Similarity 68.3%; Pred. No. 5e-37;
 Matches 316; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

Qy	93	GCCTTTTCCCTGGGGCGCTGGTGGCGGTGACCGCTGTGCTGTGCTGTGCTGTGCTG	152
Db	95	GCTTCTTCCCGCGCGCTGTGCTGGGGCGTCAAGCACCTGCTGCTGCTGCTGCTG	154
Qy	153	CGGGGTGAGCGGCAACGTGTGACCGTGTGATGCTGATCGGGCGCTACCGGGAATGCGGAC	212
Db	155	GGGTATCGTGGCAACCTGCTCACCATGCTGTGTGTGTGCTGCGCTTCCGCGAGCTGCGCAC	214
Qy	213	CACACCACTTGTACTTGGGAGCATGGCGGTGTCGGACCTACTCATCTGCTCGGGCT	272
Db	215	CACACCACTTGTACTTGGGAGCATGGCGGTGTCGGACCTACTCATCTGCTCGGGCT	274
Qy	273	GCCTTTCGACCTGTACCGCTCTGGCGCTCGCGCGCTGCTGGGTGTGCGGCGCTGCTG	332
Db	275	GCCTTTCGACCTGTGCTGCTCTGGAGTACCGGCGCTGGAATCTTGGGCGACCTCTCTG	334
Qy	333	CCGCTGTCTTCTTACGTGGGCGAGGGCTGACCTACGCGCACGCTGTGTCACATGACCGC	392
Db	335	CAAACTCTTCCAATTGCTCAGTGAGAGTGCACCTACGCGCACGCTGTCTACCATCACAGC	394
Qy	393	GCTCAGGTCGAGGCTACTTGGCCATCTGCGCGCGCTCGCGCGCGCTCTTGTGTCAC	452
Db	395	GCTGAGGTCGAGGCTACTTGGCCATCTGCTTCCCACTCCGCGCGCAAGGTGGTGTGTCAC	454
Qy	453	CCGGCGCGCGGTCGCGCGCTCATCGTGTGCTTGGGCGCTGCGCGCTGCTCTCTGCGCG	512
Db	455	CAAGGGCGGGTGAAGCTGTGTCTGTCATCTGGGCGCTGCGCGCTCTTGTGCGCGCGG	514
Qy	513	TCCTTCTTGTTCCTGTTGGGCGTGCAGCAGGACCCCGGCA	555
Db	515	GCCCATCTTCTGCTAGTGGGGTGGAGCAGGACGAGACCGGCA	557

Search completed: January 1, 2004, 02:55:55
Job time : 92.6584 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 1, 2004, 06:19:59 ; Search time 614.08 Seconds
(without alignments)
6983.649 Million cell updates/sec

Title: US-09-719-485-2

Perfect score: 1239

Sequence: 1 atggcgagcccttggaacgg.....acgtgaagacgatgggataa 1239

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2263443 seqs, 1730637950 residues

Total number of hits satisfying chosen parameters: 4526886

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/prodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/prodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/prodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/prodata/2/pubpna/US09A_PUBCOMB.seq:*
- 14: /cgn2_6/prodata/2/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/prodata/2/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1239	100.0	1239	13	US-10-206-677-1 Sequence 1, Appli
2	1239	100.0	1239	15	US-10-225-567A-472 Sequence 472, App
3	1239	100.0	1239	15	US-10-290-078-13 Sequence 13, Appl
4	1239	100.0	1239	15	US-10-290-078-14 Sequence 14, Appl
5	896.2	72.3	2040	12	US-10-417-820A-129 Sequence 129, App
6	891.4	71.9	2040	12	US-10-417-820A-151 Sequence 151, App
7	234.8	19.0	1063	13	US-10-303-204A-1 Sequence 1, Appli
8	230	18.6	1029	13	US-10-303-204A-4 Sequence 4, Appli
9	229.8	18.5	1092	13	US-10-303-204A-15 Sequence 15, Appl
10	228.8	18.5	870	15	US-10-225-567A-139 Sequence 139, App
11	228.8	18.5	1122	13	US-10-303-204A-9 Sequence 9, Appli
12	228.4	18.4	3129	13	US-10-303-204A-14 Sequence 14, Appl
13	227.8	18.4	1088	13	US-10-303-204A-6 Sequence 6, Appli
14	227.8	18.4	1101	13	US-10-276-392-22 Sequence 22, Appl
15	227.8	18.4	1101	15	US-10-251-385-87 Sequence 87, Appl

16	227.8	18.4	1101	15	US-10-251-385-209 Sequence 209, App
17	153	12.3	836	13	US-10-303-204A-11 Sequence 11, Appl
18	134.8	10.9	801	11	US-09-782-974C-17 Sequence 17, Appl
19	134.8	10.9	1239	15	US-10-225-567A-556 Sequence 556, App
20	134.8	10.9	1248	12	US-10-417-820A-11 Sequence 11, Appl
21	134.8	10.9	1248	13	US-10-272-983-11 Sequence 11, Appl
22	134.8	10.9	1248	13	US-10-393-807-11 Sequence 11, Appl
23	132	10.7	1569	15	US-10-225-567A-431 Sequence 431, App
24	127.2	10.3	1233	13	US-09-826-509-536 Sequence 536, App
25	123	9.9	1257	13	US-09-826-509-534 Sequence 534, App
26	123	9.9	4131	13	US-10-101-510-417 Sequence 417, App
27	123	9.9	4131	13	US-10-101-510-752 Sequence 752, App
28	123	9.9	4131	15	US-10-225-567A-206 Sequence 206, App
29	123	9.9	4144	13	US-10-115-831-149 Sequence 149, App
30	122.8	9.9	1529	13	US-10-205-219-22 Sequence 22, Appl
31	122	9.8	1095	15	US-09-826-509-576 Sequence 576, App
32	122	9.8	1095	15	US-10-225-567A-319 Sequence 319, App
33	122	9.8	1285	15	US-10-265-872-2 Sequence 2, Appli
34	120	9.7	1349	12	US-10-240-145-96 Sequence 96, Appl
35	120	9.7	1535	12	US-10-240-145-10 Sequence 10, Appl
36	120	9.7	1535	15	US-10-146-419-12 Sequence 12, Appl
37	120	9.7	1535	15	US-10-146-123-12 Sequence 12, Appl
38	119	9.6	1212	13	US-10-353-690-9 Sequence 9, Appli
39	119	9.6	1212	15	US-10-083-168-13 Sequence 13, Appl
40	119	9.6	1212	15	US-10-083-168-82 Sequence 82, Appl
41	119	9.6	1212	15	US-10-251-385-113 Sequence 113, App
42	119	9.6	1212	15	US-10-251-385-223 Sequence 223, App
43	119	9.6	1212	15	US-10-225-567A-539 Sequence 539, App
44	119	9.6	1212	15	US-10-290-078-16 Sequence 16, Appl
45	119	9.6	1212	15	US-10-290-078-17 Sequence 17, Appl

ALIGNMENTS

RESULT 1

US-10-206-677-1
; Sequence 1, Application US/10206677
; Publication No. US20030186336A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences, Inc.
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burmer, Glenn C.
; APPLICANT: Roush, Christine L.
; APPLICANT: Kulanter, Bruce G.
; TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC COMPOSITIONS AND METHODS RELATED TO
; TITLE OF INVENTION: GPCR 38, A G PROTEIN-COUPLED RECEPTOR (GPCR)
; FILE REFERENCE: 1920-1-8
; CURRENT APPLICATION NUMBER: US/10/206,677
; PRIOR FILING DATE: 2002-11-06
; PRIOR APPLICATION NUMBER: 60/250,251
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 60/250,452
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1239
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-206-677-1

Query Match	100.0%	Score 1239;	DB 13;	Length 1239;
Best Local Similarity	100.0%	Pred. No. 2.5e-293;		
Matches 1239;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	ATGGGAGCCCTTGGAAACGACGACGCGCCCGAGGGGCGGGAGCGCCGCTGTGCCCC	60	
Db	1	ATGGGAGCCCTTGGAAACGACGACGCGCCCGAGGGGCGGGAGCGCCGCTGTGCCCC	60	
Qy	61	GCCTGCGCGCTTGGCAACGACGCGCCGCTGTGCCCCCTTTCCCTGTGGGGGCGCTGTGCGG	120	
Db	61	GCCTGCGCGCTTGGCAACGACGCGCCGCTGTGCCCCCTTTCCCTGTGGGGGCGCTGTGCGG	120	


```
Db 601 CCTCTGCCCTGTCGCGCGCTCTCTGCGCTCTGCGGCGCGACCGCGCTCCCCGCGCTCG 660
Qy 661 GGGCCCGAGACCGCGGAGCGCGCGCTGTTAGCGCGCGAAATGCCCGCGCGAGCCCCGCG 720
Db 661 GGGCCCGAGACCGCGGAGCGCGCGCTGTTAGCGCGCGAAATGCCCGCGCGAGCCCCGCG 720
Qy 721 CAGCTGGCGCGCTGCGTGTGTCATGCTGTGGGTACCAACCGCTACTTCTTCTGCGCCTTT 780
Db 721 CAGCTGGCGCGCTGCGTGTGTCATGCTGTGGGTACCAACCGCTACTTCTTCTGCGCCTTT 780
Qy 781 CTGTGCTCAGCATCTCTACGGGCTCATCGGGGAGAGTGTGGAGCAGCGCGCGCG 840
Db 781 CTGTGCTCAGCATCTCTACGGGCTCATCGGGGAGAGTGTGGAGCAGCGCGCGCG 840
Qy 841 CTGGAGCGCGCGCGCTCGGGGCGGAGAGGCGCACCGGAGAGCCGTCGCGTCTG 900
Db 841 CTGGAGCGCGCGCGCTCGGGGCGGAGAGGCGCACCGGAGAGCCGTCGCGTCTG 900
Qy 901 CTGGTGTGTTCTGCGCATTTATAATTTGCTGTGTTGCCCTTCCAGCTTGGCAGAACTT 960
Db 901 CTGGTGTGTTCTGCGCATTTATAATTTGCTGTGTTGCCCTTCCAGCTTGGCAGAACTT 960
Qy 961 TACATAACACGAGATTCGCGGATGATGATCTCTCAGTACTTTAACATCGTCGCT 1020
Db 961 TACATAACACGAGATTCGCGGATGATGATCTCTCAGTACTTTAACATCGTCGCT 1020
Qy 1021 CTGCAACTTTTCTATCTGAGCGCATCTATCAACCAATCTCTACAACTCATTTCAAAG 1080
Db 1021 CTGCAACTTTTCTATCTGAGCGCATCTATCAACCAATCTCTACAACTCATTTCAAAG 1080
Qy 1081 AAGTACAGAGCGCGCTTTAAACTGCTGTCGCAAGGAAGTCCAGGCCGAGAGGCTTC 1140
Db 1081 AAGTACAGAGCGCGCTTTAAACTGCTGTCGCAAGGAAGTCCAGGCCGAGAGGCTTC 1140
Qy 1141 CACAGAGCAGGACACTCGGGGGAAGTTCAGGGGACACTGGAGGAGACACGCTGGC 1200
Db 1141 CACAGAGCAGGACACTCGGGGGAAGTTCAGGGGACACTGGAGGAGACACGCTGGC 1200
Qy 1201 TACACCGAGACAAGCGCTTAAGTCAAGACGATGGATAA 1239
Db 1201 TACACCGAGACAAGCGCTTAAGTCAAGACGATGGATAA 1239

RESULT 3
US-10-290-078-13
; Sequence 13, Application US/10290078
; Publication No. US20030124596A1
; GENERAL INFORMATION:
; APPLICANT: Carroll, Joseph A.
; TITLE OF INVENTION: Methods and Compositions for Treating
; TITLE OF INVENTION: Hematological Disorders Using 232, 2059, 10630, 12848, 13875,
; TITLE OF INVENTION: 14395, 14618, 17692 or 58874
; FILE REFERENCE: MF12001-288P1(M)
; CURRENT APPLICATION NUMBER: US/10/290,078
; CURRENT FILING DATE: 2002-11-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 1239
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-290-078-13

Query Match 100.0%; Score 1239; DB 15; Length 1239;
Best Local Similarity 100.0%; Pred. No. 2.5e-293;
Matches 1239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGGAGCGCCCTGGAACGGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
Db 1 ATGGGAGCGCCCTGGAACGGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
Qy 61 GCGCTGCGCGCTTCGACGAGCGCGCTGCTCGGCCCTTTTCCCTTGGGGCGCGCTGGTGGC 120
```

```
Db 61 GCGCTGCGCGCTTCGACAGAGCGCGCTGCTCGCCCTTTTCCCTTGGGGCGCGCTGGTGGC 120
Qy 121 GTGACCGCTGTGTGTCCTGTGTCCTGTGTCGTCGCGGTGAGCGGCAACGTGGTACCGGTG 180
Db 121 GTGACCGCTGTGTGTCCTGTGTCCTGTGTCGTCGCGGTGAGCGGCAACGTGGTACCGGTG 180
Qy 181 ATGCTGATCGGGCGCTTACCGGGACATGCGGACCAACCACTTGTGTACCTTGGGAGCATG 240
Db 181 ATGCTGATCGGGCGCTTACCGGGACATGCGGACCAACCACTTGTGTACCTTGGGAGCATG 240
Qy 241 GCGGTGTCGAGCATCTCTACGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 300
Db 241 GCGGTGTCGAGCATCTCTACGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 300
Qy 301 TCGCGGCGCTTGGGTGTTTGGGGCGCTGCTTGGCGCGCTGCTTGGCGCGCTGCTTGGCGCG 360
Db 301 TCGCGGCGCTTGGGTGTTTGGGGCGCTGCTTGGCGCGCTGCTTGGCGCGCTGCTTGGCGCG 360
Qy 361 TGCACCTACGCGACGCTGTGTCATGACCGCGCTCAGCGCTCAGCGCTCAGCGCTCAGCGCT 420
Db 361 TGCACCTACGCGACGCTGTGTCATGACCGCGCTCAGCGCTCAGCGCTCAGCGCTCAGCGCT 420
Qy 421 TCGCGGCGCTTGGGTGTTTGGGGCGCTGCTTGGCGCGCTGCTTGGCGCGCTGCTTGGCGCG 480
Db 421 TCGCGGCGCTTGGGTGTTTGGGGCGCTGCTTGGCGCGCTGCTTGGCGCGCTGCTTGGCGCG 480
Qy 481 GTGCTCTGGGCGCTGCGGCTGCTTGGCGCGCTGCTTGGCGCGCTGCTTGGCGCGCTGCGAG 540
Db 481 GTGCTCTGGGCGCTGCGGCTGCTTGGCGCGCTGCTTGGCGCGCTGCTTGGCGCGCTGCGAG 540
Qy 541 CAGGACCCCGGACATCTCCGTAAGTCCCGGCGCTCAATGGCACCGCGCGGATGCGCTCTCTCG 600
Db 541 CAGGACCCCGGACATCTCCGTAAGTCCCGGCGCTCAATGGCACCGCGCGGATGCGCTCTCTCG 600
Qy 601 CCTCTGCGCTGTCGCGCGCTCTGCGGCGCGCTGCGGCGCGCGCGCGCGCGCGCGCGCG 660
Db 601 CCTCTGCGCTGTCGCGCGCTCTGCGGCGCGCTGCGGCGCGCGCGCGCGCGCGCGCGCG 660
Qy 661 GGGCCCGAGACCGCGGAGCGCGCGCTGTTAGCGCGCGAAATGCCCGCGCGAGCGCGCGCG 720
Db 661 GGGCCCGAGACCGCGGAGCGCGCGCTGTTAGCGCGCGAAATGCCCGCGCGAGCGCGCGCG 720
Qy 721 CAGCTGGCGCGCTGCGTGTGTCATGCTGTGGGTACCAACCGCTACTTCTTCTGCGCCTTT 780
Db 721 CAGCTGGCGCGCTGCGTGTGTCATGCTGTGGGTACCAACCGCTACTTCTTCTGCGCCTTT 780
Qy 781 CTGTGCTCAGCATCTCTACGGGCTCATCGGGGAGAGTGTGGAGCAGCGCGCGCGCGCG 840
Db 781 CTGTGCTCAGCATCTCTACGGGCTCATCGGGGAGAGTGTGGAGCAGCGCGCGCGCGCG 840
Qy 841 CTGGAGCGCGCGCGCTCGGGGCGGAGAGGCGCACCGGAGAGCCGTCGCGTCTG 900
Db 841 CTGGAGCGCGCGCGCTCGGGGCGGAGAGGCGCACCGGAGAGCCGTCGCGTCTG 900
Qy 901 CTGGTGTGTTCTGCGCATTTATAATTTGCTGTGTTGCCCTTCCAGCTTGGCAGAACTT 960
Db 901 CTGGTGTGTTCTGCGCATTTATAATTTGCTGTGTTGCCCTTCCAGCTTGGCAGAACTT 960
Qy 961 TACATAACACGGAAGATTCGCGGATGATGATCTCTCAGTACTTTAACATCGTCGCT 1020
Db 961 TACATAACACGGAAGATTCGCGGATGATGATCTCTCAGTACTTTAACATCGTCGCT 1020
Qy 1021 CTGCAACTTTTCTATCTGAGCGCATCTATCAACCAATCTCTACAACTCATTTCAAAG 1080
Db 1021 CTGCAACTTTTCTATCTGAGCGCATCTATCAACCAATCTCTACAACTCATTTCAAAG 1080
Qy 1081 AAGTACAGAGCGCGCGCTTTAAACTGCTGTCGCAAGGAAGTCCAGGCCGAGAGGCTTC 1140
Db 1081 AAGTACAGAGCGCGCGCTTTAAACTGCTGTCGCAAGGAAGTCCAGGCCGAGAGGCTTC 1140
Qy 1141 CACAGAGCAGGAGACACTCGGGGGAAGTTCAGGGGACACTGGAGGAGACACGCTGGC 1200
```

Db 1141 CACAAAGCAGGACACTCGGGGGAAGTTGACGGGACACTGAGGAGACACGGTGGGC 1200
Qy 1201 TACACCGAGCAAGCGCTAAGCTGAAGACGATGGATAA 1239
Db 1201 TACACCGAGCAAGCGCTAAGCTGAAGACGATGGATAA 1239

RESULT 4
US-10-290-078-14
; Sequence 14, Application US/10290078
; Publication No. US20030124596A1
; GENERAL INFORMATION:
; APPLICANT: Carroll, Joseph A.
; TITLE OF INVENTION: Methods and Compositions for Treating
; TITLE OF INVENTION: Hematological Disorders Using 232, 2059, 10630, 12848, 13875,
; TITLE OF INVENTION: 14395, 14618, 17892 or 58874
; FILE REFERENCE: MPI2001-288P1(M)
; CURRENT APPLICATION NUMBER: US/10/290,078
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 1239
; TYPE: DNA
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)....(1239)
US-10-290-078-14

Query Match 100.0%; Score 1239; DB 15; Length 1239;
Best Local Similarity 100.0%; Pred. No. 2.5e-293;
Matches 1239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGGACAGCCCTGGAAACGGCAGCGACGGCCCCAGGGGGCGCGGAGCCCGGTGCC 60
Db 1 ATGGGACAGCCCTGGAAACGGCAGCGACGGCCCCAGGGGGCGCGGAGCCCGGTGCC 60

Qy 61 GCGTGGCGGCTTGGACGAGCGCGCTGTCGCGCCCTTCCCTGGGGGGCGTGGTCCG 120
Db 61 GCGTGGCGGCTTGGACGAGCGCGCTGTCGCGCCCTTCCCTGGGGGGCGTGGTCCG 120

Qy 121 GTGACCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 121 GTGACCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180

Qy 181 ATGCTGATCGGGGCGCTACCGGACATCGGACACACCACTTGTACCTGGGCGAGCATG 240
Db 181 ATGCTGATCGGGGCGCTACCGGACATCGGACACACCACTTGTACCTGGGCGAGCATG 240

Qy 241 GCGGTGTCGACCTACTCATCTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 300
Db 241 GCGGTGTCGACCTACTCATCTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 300

Qy 301 TCGGGGCGCTGGGGTGTGGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Db 301 TCGGGGCGCTGGGGTGTGGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360

Qy 361 TGCACTACCGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Db 361 TGCACTACCGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420

Qy 421 TCGGGGCGCTGCGGGCGCGGCTTGTGTACCGGGCGCGGCTGCGGGCGCTCATCGCT 480
Db 421 TCGGGGCGCTGCGGGCGCGGCTTGTGTACCGGGCGCGGCTGCGGGCGCTCATCGCT 480

Qy 481 GTGCTCTGGGCGGTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 481 GTGCTCTGGGCGGTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540

Qy 541 CAGACCCCGGACATCTCCGTAGTCCCGGGGCTCAATGGACCGCGCGGATCGGCTCTCG 600
Db 541 CAGACCCCGGACATCTCCGTAGTCCCGGGGCTCAATGGACCGCGCGGATCGGCTCTCG 600

Qy 601 CCTCTCGCCTCGTCCGCGCCTCTCTGGCTCTCGCGGGCGCCACGCGCGTCCCCCGCTCG 660
Db 601 CCTCTCGCCTCGTCCGCGCCTCTCTGGCTCTCGCGGGCGCCACGCGCGTCCCCCGCTCG 660

Qy 661 GGGCCCGAGACCGCGGAGCGCGCGCTCTTTAGCGCGGAATGCGCGGCGAGCCCGCG 720
Db 661 GGGCCCGAGACCGCGGAGCGCGCGCTCTTTAGCGCGGAATGCGCGGCGAGCCCGCG 720

Qy 721 CAGCTGGCGCGCTGCGTGTATGCTGTGGGTACACCGCCTACTTCTTCTGCTGCCCTTT 780
Db 721 CAGCTGGCGCGCTGCGTGTATGCTGTGGGTACACCGCCTACTTCTTCTGCTGCCCTTT 780

Qy 781 CTGTGCTCAGCATCTCTTACGGGCTCATCGGGCGGAGTGTGGAGACGCGCGGCGCG 840
Db 781 CTGTGCTCAGCATCTCTTACGGGCTCATCGGGCGGAGTGTGGAGACGCGCGGCGCG 840

Qy 841 CTGCGAGGCGCGCGCGCTCTCGGGCGGAGAGAGCCACCGGCGAGACCGTCCGCGTCTG 900
Db 841 CTGCGAGGCGCGCGCGCTCTCGGGCGGAGAGAGCCACCGGCGAGACCGTCCGCGTCTG 900

Qy 901 CTGGTGGTGTCTTCTGCGATTTATAATTTGCTGGTGGCTTCCACGTTGGAGAAATCAT 960
Db 901 CTGGTGGTGTCTTCTGCGATTTATAATTTGCTGGTGGCTTCCACGTTGGAGAAATCAT 960

Qy 961 TACATAAACACGGAAGATTTCGCGGATGATGCTTCTCTCAGTACTTTAAACATCGTCGT 1020
Db 961 TACATAAACACGGAAGATTTCGCGGATGATGCTTCTCTCAGTACTTTAAACATCGTCGT 1020

Qy 1021 CTGCAACTTTTCTATCTGAGCGCATCTATCAACCAATCTCTTACAACTCATTTCAAAG 1080
Db 1021 CTGCAACTTTTCTATCTGAGCGCATCTATCAACCAATCTCTTACAACTCATTTCAAAG 1080

Qy 1081 AAGTACAGAGCGCGCGCTTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
Db 1081 AAGTACAGAGCGCGCGCTTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140

Qy 1141 CACAGAAGCAGGAGACACTGCGGGGAGTTGCGGGGACACTGAGGAGACACGCTGGGC 1200
Db 1141 CACAGAAGCAGGAGACACTGCGGGGAGTTGCGGGGAGTTGAGGAGACACTGAGGAGACACGCTGGGC 1200

Qy 1201 TACACCGAGCAAGCGCTAACGTCAGACGATGGGATAA 1239
Db 1201 TACACCGAGCAAGCGCTAACGTCAGACGATGGGATAA 1239

RESULT 5
US-10-417-820A-129
; Sequence 129, Application US/10417820A
; Publication No. US20030229216A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Ruoping
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lowitz, Kevin
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Behan, Dominic P.
; TITLE OF INVENTION: Constitutively Activated Human G Protein Coupled
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: 7. US28. CON
; CURRENT APPLICATION NUMBER: US/10/417,820A
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: 09/416,760
; PRIOR FILING DATE: 1999-10-12
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: 60/110,060
; PRIOR FILING DATE: 1998-11-27
; PRIOR APPLICATION NUMBER: 60/120,416
; PRIOR FILING DATE: 1999-02-16
; PRIOR APPLICATION NUMBER: 60/121,852
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 60/109,213
; PRIOR FILING DATE: 1998-11-20

```

; PRIOR APPLICATION NUMBER: 60/123,944.
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/123,945
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/123,948
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/123,951
; PRIOR FILING DATE: 1999-03-12
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 129
; LENGTH: 2040
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-417-820A-129

Query Match      72.3%; Score 896.2; DB 12; Length 2040;
Best Local Similarity 99.7%; Pred. No. 1.8e-209;
Matches 898; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATGGGCGAGCCCTTGGACGCGAGCGCGCCCGGAGGGGGCGGAGCGCGCGTGGGCC 60
Db 1 ATGGGCGAGCCCTTGGACGCGAGCGCGCCCGGAGGGGGCGGAGCGCGCGTGGGCC 60
Qy 61 GCGCTGCGCGCTTGGCGAGCGCGCGTGTCTGCGCCCTTTCCCTTGGGGCGCTGGTGCCG 120
Db 61 GCGCTGCGCGCTTGGCGAGCGCGCGTGTCTGCGCCCTTTCCCTTGGGGCGCTGGTGCCG 120
Qy 121 GTGACCGCTGTGTGCTGTGCTGTCTGCTGCGGGTGTGAGCGGCAAGTGTGACCGTG 180
Db 121 GTGACCGCTGTGTGCTGTGCTGTCTGCTGCGGGTGTGAGCGGCAAGTGTGACCGTG 180
Qy 181 ATGTGTATCGGGCGCTACCGGAGCATGGGACACACCACTTGTAACCTGGGCGAGCATG 240
Db 181 ATGTGTATCGGGCGCTACCGGAGCATGGGAGACACCACTTGTAACCTGGGCGAGCATG 240
Qy 241 GCCGTGTCCGACCTACTCATCTGCTCGGGGTGCGCTTCGACCTGTACCGCTCTTGGCGC 300
Db 241 GCCGTGTCCGACCTACTCATCTGCTCGGGGTGCGCTTCGACCTGTACCGCTCTTGGCGC 300
Qy 301 TCGGGCGCTGGGTGTTGGGGCGGTGTCTGCTGCGCGCTGTCCCTCTAGTGGGCGAGGGC 360
Db 301 TCGGGCGCTGGGTGTTGGGGCGGTGTCTGCTGCGCGCTGTCCCTCTAGTGGGCGAGGGC 360
Qy 361 TGCACCTACGCGACGCTGCTGCATGACCGGCTCAGCGTCGAGCGCTACTCGCCATC 420
Db 361 TGCACCTACGCGACGCTGCTGCATGACCGGCTCAGCGTCGAGCGCTACTCGCCATC 420
Qy 421 TCGCGCGCGCTCCGCGCGCGGTCTTGGTCAACCGCGCGCGCTCCGCGCGCTCATCGCT 480
Db 421 TCGCGCGCGCTCCGCGCGCGGTCTTGGTCAACCGCGCGCGCTCCGCGCGCTCATCGCT 480
Qy 481 GTGCTCTGGGCGGTGGCGCTGCTCTGCTGCGGGTCCCTTTCTTGTCTGGTGGGCGTCAAG 540
Db 481 GTGCTCTGGGCGGTGGCGCTGCTCTGCTGCGGGTCCCTTTCTTGTCTGGTGGGCGTCAAG 540
Qy 541 CAGGACCCCGGCANTCCGTAGTCCCGGGCTCAATGGCACCGCGCGCATCGCTCCCTCG 600
Db 541 CAGGACCCCGGCANTCCGTAGTCCCGGGCTCAATGGCACCGCGCGCATCGCTCCCTCG 600
Qy 601 CCTCTCGCTCTGCTCGCGCGCTCTTGGGTCTCGGGGGCGCACCGCGCTCCCGCGCGTGC 660
Db 601 CCTCTCGCTCTGCTCGCGCGCTCTTGGGTCTCGGGGGCGCACCGCGCTCCCGCGCGTGC 660
Qy 661 GGGCGCGAGACCGGGAGGCGCGCGCTGTTCAGCCCGGAATACCGGCGCAGCCCGCGC 720
Db 661 GGGCGCGAGACCGGGAGGCGCGCGCTGTTCAGCCCGGAATACCGGCGCAGCCCGCGC 720
Qy 721 CAGCTGGGCGCGCTGCGGTGTCATGCTGTGGGTCAACCGCGCTACTTCTTCTGCGCTTT 780
Db 721 CAGCTGGGCGCGCTGCGGTGTCATGCTGTGGGTCAACCGCGCTACTTCTTCTGCGCTTT 780

; Sequence 151, Application US/10417820A
; Publication No. US20030229216A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Ruoping
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lowitz, Kevin
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Behan, Dominic P.
; TITLE OF INVENTION: Constitutively Activated Human G Protein Coupled
; FILE REFERENCE: 7.US28.CON
; CURRENT APPLICATION NUMBER: US/10417,820A
; PRIOR FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: 09/416,760
; PRIOR FILING DATE: 1999-10-12
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: 60/110,060
; PRIOR FILING DATE: 1998-11-27
; PRIOR APPLICATION NUMBER: 60/120,416
; PRIOR FILING DATE: 1999-02-16
; PRIOR APPLICATION NUMBER: 60/121,852
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 60/109,213
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/123,944
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/123,945
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/123,948
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/123,951
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 151
; LENGTH: 2040
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-417-820A-151

Query Match      71.9%; Score 891.4; DB 12; Length 2040;
Best Local Similarity 98.8%; Pred. No. 2.7e-208;
Matches 898; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 1 ATGGGCGAGCCCTTGGAAACGCGAGCGCGCCCGGAGGGGGCGCGGAGCGCGCGTGGGCC 60
Db 1 ATGGGCGAGCCCTTGGAAACGCGAGCGCGCGCCCGGAGGGGGCGCGGAGCGCGCGTGGGCC 60
Qy 61 GCGCTCGCGCTTGGCGAGCGCGCGTGTCTGCGCCCTTTCCCTTGGGGCGCTGGTGCCG 120
Db 61 GCGCTCGCGCTTGGCGAGCGCGCGTGTCTGCGCCCTTTCCCTTGGGGCGCTGGTGCCG 120
Qy 121 GTGACCGCTGTGTGCTGTGCTGTCTGCTGCGGGTGTGAGCGGCAAGTGTGACCGTG 180
Db 121 GTGACCGCTGTGTGCTGTGCTGTCTGCTGCGGGTGTGAGCGGCAAGTGTGACCGTG 180
```

```

Qy 181 ATGCTGATCGGGCGCTACCGGAGCATCGGACCATCGGACCACTTGTACCTGGGAGCATG 240
Db 181 ATGCTGATCGGGCGCTACCGGAGCATCGGACCATCGGACCACTTGTACCTGGGAGCATG 240
Qy 241 GCGGTGTCGACCTACTCATCTGCTGGGCTGCGGCTTGCAGCTGTACCGCTTGGGCG 300
Db 241 GCGGTGTCGACCTACTCATCTGCTGGGCTGCGGCTTGCAGCTGTACCGCTTGGGCG 300
Qy 301 TCGGGCGCTGGGTGTTGGGCGCGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 360
Db 301 TCGGGCGCTGGGTGTTGGGCGCGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 360
Qy 361 TGCACCTACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Db 361 TGCACCTACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Qy 421 TCGCGCGCTGCTGCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 421 TCGCGCGCTGCTGCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Qy 481 GTGCTGCTGGGCGCTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 481 GTGCTGCTGGGCGCTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Qy 541 CAGGACCCCGGCTATCCGCTAGTCCGCGGCTCAATGCGACCGCGGATCGCTCTCTCG 600
Db 541 CAGGACCCCGGCTATCCGCTAGTCCGCGGCTCAATGCGACCGCGGATCGCTCTCTCG 600
Qy 601 CCTCTGCTGCTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 601 CCTCTGCTGCTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Qy 661 GGGCCGAGACCGGCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Db 661 GGGCCGAGACCGGCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Qy 721 CAGTGGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Db 721 CAGTGGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Qy 781 CTGTGCTGAGCTCTCTACGCGCTCATCGGCGCTCATCGGCGGAGCTGTGGAGCGCGCG 840
Db 781 CTGTGCTGAGCTCTCTACGCGCTCATCGGCGCTCATCGGCGGAGCTGTGGAGCGCGCG 840
Qy 841 CTGGAGCGCGCGCTGCTGCGGCGGAGAGGCGGCGGAGGCGGCGGAGGCGGCGGAG 900
Db 841 CTGGAGCGCGCGCTGCTGCGGCGGAGAGGCGGCGGAGGCGGCGGAGGCGGCGGAG 900
Qy 901 CTGGTGGTG 909
Db 901 CGTAAGTTG 909

```

```

RESULT 7
US-10-303-204A-1
; Sequence 1, Application US/10303204A
; Publication No. US20030166144A1
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feigner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus H. T.
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR
; FILE REFERENCE: 19589PCA
; CURRENT APPLICATION NUMBER: US/10/303,204A
; PRIOR FILING DATE: 2002-11-25
; PRIOR APPLICATION NUMBER: 09/077,674
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: PCT/US96/19445

```

```

; PRIOR FILING DATE: 1996-12-10
; PRIOR APPLICATION NUMBER: 60/018,962
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: 60/008,582
; PRIOR FILING DATE: 1995-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1063
; TYPE: DNA
; ORGANISM: sus scrofa
US-10-303-204A-1

Query Match 19.0%; Score 234.8; DB 13; Length 1063;
Best Local Similarity 69.3%; Pred No. 8.8e-48;
Matches 320; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

Qy 92 CGCCCTTTCCCTTGGGGCGCTGTCGCGGTGACCGCTGTGTGCTGTGCTGTGCTGTGCTG 151
Db 69 CGCTCTTTCCCAACGCGCTGTTGGCGGGCGTCAACCGCACCTGCTGCTGCTGCTGCTG 128
Qy 152 TCGGGGTGAGCGGCAACGTTGATCGGTGATGCTGATCGGGCGCTACCGGAGATCGGGA 211
Db 129 TGGGTATCGCGGGCAACCTGCTCACGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 188
Qy 212 CCACCACTTGTATCTGCGGAGCATGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 271
Db 189 CCACCACTTGTATCTGCGGAGCATGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 248
Qy 272 TCGGCTTGCAGCTGTACCGCTTGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCT 331
Db 249 TCGCCCTGAGCTTTCGCGCTTGGCAGTACCGGCTTGGAACTTGGCAACTGCTCT 308
Qy 332 GCGGCTGCTGCTGCTGCTGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 391
Db 309 GCAAACTCTTCCAGTTCTGTTAGCGAGAGTGCACCTACGCGCACGTGCTCACCATCAC 368
Qy 392 CGCTAGCGTCGAGCGCTACCTGCGCATCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCT 451
Db 369 CGCTAGCGTCGAGCGCTACCTGCGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 428
Qy 452 CCGGCGCGCGCTGCGCGCGCTCATCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 511
Db 429 CCAAGCGCGGTTAAAGCTGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 488
Qy 512 GTCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 553
Db 489 GGCCCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 530

```

```

RESULT 8
US-10-303-204A-4
; Sequence 4, Application US/10303204A
; Publication No. US20030166144A1
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feigner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus H. T.
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR
; FILE REFERENCE: 19589PCA
; CURRENT APPLICATION NUMBER: US/10/303,204A
; PRIOR FILING DATE: 2002-11-25
; PRIOR APPLICATION NUMBER: 09/077,674
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: PCT/US96/19445
; PRIOR FILING DATE: 1996-12-10
; PRIOR APPLICATION NUMBER: 60/018,962
; PRIOR FILING DATE: 1996-06-06

```

[illegible]

Db	600	CAAACTCTTCAATTCCTCAGTGAAGCTGCACCTACGCCACGGTGTCTACCATCAACG	659
Qy	393	GCTCAGCGTCGAGCGCTACCTGGCCATCTCCGCGCCGCTCCGCGCCCGCTCTTGGTCA	452
Db	660	GCTGAGCGTCGAGCGCTACTTGGGCATCTGCTTCCCACTCCGGGCCAAGGTGGTGGTCA	719
Qy	453	CCGGCGCGCGTCCGCGCGCTCATCGGTGTGCTCTGGGCGGTGGCGGTCTCTCGCCG	512
Db	720	CAAGGGCGGGTGAAGCTGTGCATCTTCGTATCTGGGCGGTGGCCCTTCTGCAGCGCCGG	779
Qy	513	TCCTCTTTCTTCTGTGTGGCGGTGAGCAGGACCCCGGCATCT - CGGTAGTCCCGGGCC	571
Db	780	GCCCATCTTCGTGTACTCGGGTGGAGCACGAGAACCGCACCGACTTTGGGACACCAA	839
Qy	572	TCAATGCAAGCGCGGATCGCTCTCTCGCTCTCGCTCTCGCGCCCTCTCTGGTCT	631
Db	840	CGAGTGGCGGCCACCCAGTTTGGGTGCGCTCTGGACTGCTACGGTCAATGTTGGGT	899
Qy	632	CGGGGGCCACCGCCCTCCCGCGTCTGGGGCCCGAGACCGCGGAGCGCGCGCGCTGT	691
Db	900	GTCCAGCATCTTCTTCTCTCTCTCTGTCTCAGGTCTCTACAGTCTCATCGG	959
Qy	692	TCAGCCCGGAATGCCGCGGAGCCCCCGGAGCTTGGGCGGTCTGCTGTCTATGCTGTGGG	751
Db	960	CAGGAAGCTGTGGCGGAGGAGCGCGCATGCTGTCTGGGTGCTCTCGTCAAGGACCA	1019
Qy	752	TCACCAACCGCTACTTCTCTGCGCTTTCTGTGCTCTCAGCATCTCTACGGGCTCATCG	811
Db	1020	GAACCAACGAAACCGTGAATGCTGGGTGGGTCTCAGCGCGGTCTCAGGCTTCTCT	1079
Qy	812	GGCGGGAGCTGT	823
Db	1080	CGCGGTCCTAT	1091

RESULT 12

```

US-10-303-204A-14
; Sequence 14, Application US/10303204A
; Publication No. US20030166144A1
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus H. T.
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR
; TITLE OF INVENTION: FAMILY
; FILE REFERENCE: 19589PCA
; CURRENT APPLICATION NUMBER: US/10/303,204A
; CURRENT FILING DATE: 2002-11-25
; PRIOR APPLICATION NUMBER: 09/077,674
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: PCT/US96/19445
; PRIOR FILING DATE: 1996-12-10
; PRIOR APPLICATION NUMBER: 60/018,962
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: 60/008,582
; PRIOR FILING DATE: 1995-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 3129
; TYPE: DNA
; ORGANISM: rattus norvegicus
US-10-303-204A-14

```

[illegible]

RESULT 13

```

US-10-303-204A-6
; Sequence 6, Application US/10303204A
; Publication No. US20030166144A1
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus H. T.
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR
; TITLE OF INVENTION: FAMILY
; FILE REFERENCE: 19589PCA
; CURRENT APPLICATION NUMBER: US/10/303,204A
; CURRENT FILING DATE: 2002-11-25
; PRIOR APPLICATION NUMBER: 09/077,674
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: PCT/US96/19445
; PRIOR FILING DATE: 1996-12-10
; PRIOR APPLICATION NUMBER: 60/018,962
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: 60/008,582
; PRIOR FILING DATE: 1995-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1088
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-303-204A-6

```

	Query Match	18.4%	Score 227.8;	DB 13;	Length 1088;
	Best Local Similarity	69.3%;	Pred. No. 4.5e-46;		
	Matches 316; Conservative	0;	Mismatches 147;	Indels 0;	Gaps 0;
Qy	93	GCCCTTTCCTGGGCGCCTGTGTGCCGTGCACCGCTGTGCTGTGCCCTGTGCCTGTTCGTTCGTCGT	152		
Dd	95	GCTCTTCCCAGCGCCGCTGTGTGGCGGCGCTCACAGCAACCTGCGTGCGCATCTTTTCGTGGT	154		

Qy		513	TCCCTTCTTGTTCTCGTGGCGTCGAGCAGGACCCCGGCATC	555
Db		528	GCCCATTCTCGTGCTAGTCGGGTGGAGCACGAGAAGGCCACC	570

Search completed: January 1, 2004, 06:44:54
Job time : 617.08 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 31, 2003, 21:38:58 ; Search time 2904.98 Seconds
(without alignments)
10366.072 Million cell updates/sec

Title: us-09-719-485-2

Perfect score: 1239

Sequence: 1 atgggcagcccttggaagg.....acgtgaagacgatgggataa 1239

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_man:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	230	18.6	4435	11 AK049671	Mus muscu
2	226.8	18.3	500	10 BF603623	BF603623 269181 MA
3	219.2	17.7	608	14 BY724644	BY724644 BY724644
4	186.4	15.0	843	13 BU553576	BU553576 AGENCOURT

C	5	173.4	14.0	849	13	BU568940
C	7	141.6	11.4	504	28	AZ451922
	8	127	10.3	855	12	BI757121
	9	127	10.3	1010	10	BM925480
	10	125.4	10.1	634	10	BG714306
	11	125.2	10.1	1055	12	BM808829
	12	124.8	10.1	540	12	BM737726
	13	124	10.0	1027	12	BM543497
	14	122.8	9.9	1144	12	BM546496
C	15	119	9.6	941	29	CNS036WH
	16	117.8	9.5	1603	12	BM808823
	17	111.8	9.0	624	9	AU244232
	18	111.4	9.0	738	12	BI600170
	19	111.4	9.0	426	14	CB695276
	20	111.2	9.0	574	14	CB608597
	21	109.6	8.8	928	12	BI758259
	22	109.2	8.8	993	13	BO721315
	23	106.4	8.6	448	10	BF323227
	24	105.6	8.5	3277	11	AK036756
	25	105	8.5	2760	11	AK081073
	26	104.8	8.5	427	13	BY284735
	27	104.6	8.4	828	12	BI597845
	28	103.6	8.4	660	14	BY728041
	29	102.6	8.3	866	14	CD246184
	30	102.4	8.3	980	12	BM543468
	31	102.2	8.2	419	13	BY273314
	32	102.2	8.2	994	14	BY705540
	33	99	8.2	1189	11	AK005368
	34	98.8	8.0	397	14	CB772331
	35	98.4	7.9	1918	11	AK053776
C	36	98.4	7.9	751	14	CB154463
	37	97	7.8	1233	14	CA975828
	38	97	7.8	785	12	BI754749
C	39	96.8	7.8	843	29	CNS04857
	40	96.8	7.8	670	29	AG046172
	41	96.8	7.8	720	12	BI753905
	42	96.6	7.8	1307	11	CNSL11BD
	43	96.6	7.8	555	10	BE751626
	44	95	7.7	825	9	AU079556
	45	95	7.7	463	13	BY257194
				502	13	BX280512

ALIGNMENTS

RESULT 1

AK049671

LOCUS

DEFINITION

AK049671 4435 bp mRNA linear HTC 05-DEC-2002
Mus musculus 12 days embryo spinal cord cDNA, RIKEN full-length
enriched library, clone: C530020122 product: GROWTH HORMONE
SECRETAGUE RECEPTOR TYPE 1 homolog [Rattus norvegicus], full
insert sequence.

ACCESSION

AK049671

VERSION

AK049671.1 GI:26340405

KEYWORDS

HTC; CAP trapper.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE

1

Carninci, P., Shibata, Y., Hayata, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayaishizaki, Y.

Normalisation and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

20499374

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source

Location/Qualifiers

1. 608

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="A430103P18"

/tissue_type="thymus"

/dev_stage="0 day neonate"

/lab_hosts="DH10B"

/clone_lib="RIKEN full-length enriched, 0 day neonate thymus"

/note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'

GAGAGAGAGATCCAGAGCTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATTCGAGTTAAATTAATCCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

BASE COUNT

ORIGIN

Query Match 17.7%; Score 219.2; DB 14; Length 608;

Best Local Similarity 70.3%; Pred. No. 9.2e-36;

Matches 293; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 137 TGTGCTTGTCTGCTCGGGTGGCGGACGCTGTGACCGTGATGCTGATCGGGCGCT 196

DB 2 TGGCGCTCTTCTGGTGGGATCTCTCGGGCAACCTGTCTCACCATGCTGGTGTGTCGCGCT 61

QY 197 ACCGGGACATGCGGACACACCACTTGTACCTTGGGACGATGGCGGTGCGACCTAC 256

DB 62 TCGGGAGCTGGGACACACCACTTGTACCTTACCTATCCAGATGGCTTCTCGATCTGC 121

QY 257 TCATCTCTGCTCGGCTCGGCTTGCACCTGTACCGCTCTTGGCGCTCGCGGCTTGGGTGT 316

DB 122 TCATCTCTGCTCGGCTCGGCTTGCACCTGTACCGCTCTTGGCGCTTGGCGCTTGGAACT 181

QY 317 TCGGGCGCTGCTGCTGGGCTGTCCCTCTACGTGGGCGAGGGGTGACCTAGCCACGC 376

DB 182 TCGGCGACCTGCTCTGCAAACTCTTCCAGTTTGTTCAGCGAGAGCTGACCTAGCCACGG 241

QY 377 TGTGCAATGACCGGCTCAGGTCGAGCGCTACCTGGCCATCTGCGGCTCGCGCTCGCG 436

DB 242 TCCTACCATCAGCGGCTGAGGTCGAGGCTACTTTCGCCATCTGTTCCCGCTGCGGG 301

QY 437 CCGCGCTTGTGGTACCGGCGCGGCTCGCGGCTCATCGTGTGCTCTGGGCGGTGG 496

Db 302 CCAAGGTGGTGGTCCACCAAGGCCGTGTGAAGCTGGTCACTCTTGTCACTTTGGCGCGTGG 361

QY 497 CGCTGCTCTCTGCGGTCCTCTTGTCTCTGCTGGGCGTCGAGCAGACCCGCGCA 553

DB 362 CTTTCTGACGCGCGGCCCATCTTCTTGTCTGGTGGCGTGGAGCACGAGACGCA 418

RESULT 4

BU553576

LOCUS

DEFINITION

AGENCOURT 10242213 NIH_MGC_109 Homo sapiens cDNA clone

IMAGS:6577973 5', mRNA sequence.

ACCESSION

BU553576

VERSION

BU553576.1 GI:22903848

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 843)

AUTHORS

NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: LLC2778 row: p column: 05

High quality sequence stop: 534.

Location/Qualifiers

1. 843

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6577973"

/tissue_type="teratocarcinoma, cell line"

/lab_hosts="DH10B (phage-resistant)"

/clone_lib="NIH MGC 109"

/notes="Organ: ovary; Vector: pOTB7; Site 1: EcoRI; Site 2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC Library."

BASE COUNT 213 a 169 c 212 g 247 t 2 others

ORIGIN

Query Match 15.0%; Score 186.4; DB 13; Length 843;

Best Local Similarity 91.0%; Pred. No. 7.4e-29;

Matches 243; Conservative 0; Mismatches 17; Indels 7; Gaps 4;

QY 906 GTGGTTCTGGCAATTTATATTTGCTGGTTGCCCTTCCACGTTGGCAGAAATCAATTTACAT 965

DB 534 GGTGTTCTGGCAATTTATATTTGCTGGTTGCCCTTCCACGTTGGCAGAAATCAATTTACAT 593

QY 966 AACACGGAGATTCGGGATGATGATCTCTCTCAGTACTTTAACTCTCGCTCTGCA 1025

DB 594 AACACGGAGATTCGGGATGATGATCTCTCTCAGTACTTTAACTCTCGCTCTGCA 653

QY 1026 ACTTTTCTATCTCAGCGCATCTATCAACCAATCTCTACAACTC-ATTTCAAAGAAGT 1084

DB 654 ACTTTTCTATCTCAGCGCATCTATCAACCAATCTCTCTACAACTCAATTTCAAAGAAGT 713

QY 1085 AC-AGAGCGCGCGCTTTAAA-----CTGTGCTCGCAAGGAAGTCCAGGCGGAG-AGGCT 1138

DB 714 ACAAGAGCGCGCGCTTTAAAACACTGCTGCTTCGCAAGGGAAGTCCAGGCGGAGAGGCT 773


```

Qy 1139 TCCACAGAGCAGGACACACTCGCGGG 1165
Db 774 CCCACAGAGCAGGACACTCGCGGG 800

RESULT 5
BU568940/c
LOCUS BU568940 849 bp mRNA linear EST 16-SEP-2002
DEFINITION AGENCOURT 10400272 NIH_MGC_82 Homo sapiens cDNA clone IMAGE:6616239
5', mRNA sequence.
ACCESSION BU568940
VERSION BU568940.1 GI:22919240
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 849)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: csapbe-remail.nih.gov
Tissue Procurement: CLONTECH
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM2859 row: j column: 15
High quality sequence stop: 429.
FEATURES
Location/Qualifiers
1..849
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6616239"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_82"
/notes="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggccgcccggcc); Site_2: SfiI (ggccatcatggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGCGGCGGCAGCATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size
1.35 kb (range 0.9-4.0 kb). 14/15 colonies contained
inserts by PCR. This library was enriched for full-length
clones and was constructed by Clontech Laboratories (Palo
Alto, CA)."
BASE COUNT 199 a 229 c 193 g 228 t
ORIGIN

Query Match 14.0%; Score 173.4; DB 13; Length 849;
Best Local Similarity 88.6%; Pred. No. 3.9e-26;
Matches 233; Conservative 0; Mismatches 25; Indels 5; Gaps 4;

Qy 981 GCGGATGATGACTTCTCTAGTACTTTAA-CATCGTGGCTCTG--CAACTTTTCTATCT 1037
Db 698 GGGAAAGAGGACTTTTCTCTAGTACTTTAACTTGTGCGTTTGGCACTTTTTTTTGT 639

Qy 1038 GAGGCGATCTATCAACCAATCCCTCAACCTCATTTCAAGAGTACAGAG-CGGCGG 1096
Db 638 AGGGGCAATTTATCAACCAATCCCTTACCTTCAATTTCAAGAGTACAGAGCGCGCGG 579

Qy 1097 CTTTAAACTGCTGCTCGCAAGGAAGTCCAGGCGGAGGCTTCCACAGAGCAGGAGCA 1156
Db 578 CTTTAAACTGCTGCTCGCAAGGAAGTCCAGGCGGAGGCTTCCACAGAGCAGGAGCA 520

Qy 1157 CTGCGGGGAAGTTGTCAGGGGACACTGAGGAGACACGGTGGGCTACACCGAGACAGCG 1216
Db 519 CTGCGGGGAAGTTGTCAGGGGACACTGAGGGGACACGGTGGGCTACACCGAGACAGCG 460

```

```

Qy 1217 CTAACGTGAAGACGATCGGATAA 1239
Db 459 CTAACGTGAAGACGATCGGATAA 437

RESULT 6
AZ451922/c
LOCUS AZ451922 504 bp DNA linear GSS 04-OCT-2000
DEFINITION 1M0251C06R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0251C06 R, genomic survey sequence.
ACCESSION AZ451922
VERSION AZ451922.1 GI:10608187
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 504)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0251 row: C column: 06
Seq primer: CACACAGGAACACGCTATGACC
Class: plasmid ends
High quality sequence stop: 504.
FEATURES
Location/Qualifiers
1..504
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="UUGC1M0251C06"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/notes="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
BASE COUNT 156 a 110 c 107 g 131 t
ORIGIN

Query Match 11.4%; Score 141.6; DB 28; Length 504;
Best Local Similarity 78.4%; Pred. No. 1.6e-19;

```

```

Matches 182; Conservative 0; Mismatches 49; Indels 1; Gaps 1;
QY 899 TGCTGGTGGTGGTCTGGCATTATATAATTGTC-TGGTTGCCCTTCCACGTTGGCAGATC 957
Db 252 TGGTGGTGGAGTTCTGGCATTATAGTTTGGTTGGTTGCTTCCAAAGTTGGCAGATC 193
QY 958 ATTATACATAACACGGAAGATTCCGGGATGATGTAATCTCTCAGTACTTTAAACATCGTC 1017
Db 192 TTTTTCATGAACACTGAATAATTCGGGGTAATGCACTCTCTCGATATCTTTAATATGGTC 133
QY 1018 GCTCTGCAACTTTTCTATCTGAGCGCATCTATCAACCAATCTCTCAACCTCATTTCA 1077
Db 132 ACTCTCAACTTCTATACCGAGTGCAATCCATCAAGCCCATCTCTCCACAACCTCATGTC 73
QY 1078 AAGAAGTACAGAGCGCGGCTTTAAACTGCTGCTGCAAGGAGTCCAGGC 1129
Db 72 AATTAGTACAAAGCAGCTGTGAGGAACTGCTTCTGCGGAGACAGTCCAGGC 21

RESULT 7
LOCUS BM925480
DEFINITION 6030882F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5200868 5',
mRNA sequence.
ACCESSION BM925480
VERSION BM925480.1 GI:15748699
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM1503 row: d column: 21
High quality sequence stop: 803.
Location/Qualifiers
1..855
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5200868"
/lab_host="DH10B"
/clone_lib="NIH_MGC_114"
/notes="Organ: Brain; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 019. Note:
this is a NIH MGC Library."
BASE COUNT 103 a 300 c 299 g 153 t
ORIGIN
Query Match 10.3%; Score 127; DB 12; Length 855;
Best Local Similarity 55.2%; Pred. No. 2e-16;
Matches 294; Conservative 0; Mismatches 230; Indels 9; Gaps 2;
QY 27 CGGCCCCGAGGGCGCGGAGCGCCGCGCTGCGCGCTTGCACGAGCGCGG 86
Db 60 CAGCAGCCCGCGCGCGCGCGCGAGTCTCAACCCGCGGCTGAGCCGTGACGCCCGCGCT 119

```

```

QY 87 CTGCTGCCCCCTTTCCCTTGGGGCGCTGGTGGCGGTGACCGCTGTGTGCTGTGCTGTGCTT 146
Db 120 GGGCGTGGACACTCGCTCTGGGCCCAAGGTGCTGTTCCACCGCTCTACGCACTCATCTG 179
QY 147 CGTCTGTCGGGGTGAAGCGGCAACGTG-----GTGACCGGTGATGCTGATCGGGGCTACCG 200
Db 180 GGGCGTGGGGCGCGCGGCAATGCGCTGTCGCTGCACGTGGTCTGAAAGCGCGGCGCG 239
QY 201 GGACATGCGGACCAACCAACCACTTGTACCTGGGCGAGCATGCGGCTGTCGACCTACTCAT 260
Db 240 GCGCGGGGGCGGCTGCGGCCACCACTGCTCAGCTTGGCGCTCGCGGCGCTGCTGCTGCT 299
QY 261 CTGCTCGGGCTCGCGTTGACCTTACCGCTCTGCGCTCGCGG---CCCTGGGTGTT 317
Db 300 GCTGTTGCGGGTGGCGTGGAGCTCTACAGCTTCTGTTGTTCCACTACCCCTGGGTCTT 359
QY 318 CGGGCGCTGCTGTCGCGCTGCTCTTACGTGGGCGAGGGCTGACACTACGCGCT 377
Db 360 CGCGGACCTGGGCTGCGCGGCTACTACTTCTGTCACGAGCTGTGGCGCTTACGCGCGT 419
QY 378 GCTGCACATGACCGCTCAGCTGAGCGCTACCTGGCCATCTGCGCGCGCTCCGCGC 437
Db 420 GCTGAGCGTGGCGGCTGAGCGCGCTGAGCGCTGCTAGCGCTGTGCGCGCTTACGCGTGC 479
QY 438 CCGCGTCTTGGTCAACCGCGCGCTGCGCGGCTCATCGCTGTGCTCTGGGCGCTGGC 497
Db 480 CCGCAGCTGTGACGCGCACCGCGCGGCTGGTGGCTCTCTGTGGGCGCGCTC 539
QY 498 GCTGCTCTGCGCGTCCCTTCTTGTCTGTTGGGCGTGCAGACGAGACCCCG 550
Db 540 GCTCGGCTCGCGCTGCCATGCGCGCTCATCATGCGGCGAGAGCAGCACTCG 592

RESULT 8
LOCUS BM925480
DEFINITION 5', mRNA sequence.
ACCESSION BM925480
VERSION BM925480.1 GI:19375859
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM12814 row: n column: 04
High quality sequence stop: 703.
Location/Qualifiers
1..1010
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5763267"
/lab_host="DH10B"
/clone_lib="NIH_MGC_114"
/notes="Organ: Brain; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 019. Note:
this is a NIH MGC Library."
BASE COUNT 103 a 300 c 299 g 153 t
ORIGIN
Query Match 10.3%; Score 127; DB 12; Length 855;
Best Local Similarity 55.2%; Pred. No. 2e-16;
Matches 294; Conservative 0; Mismatches 230; Indels 9; Gaps 2;
QY 27 CGGCCCCGAGGGCGCGGAGCGCCGCGCTGCGCGCTTGCACGAGCGCGG 86
Db 60 CAGCAGCCCGCGCGCGCGCGAGTCTCAACCCGCGGCTGAGCCGTGACGCCCGCGCT 119

```

upon cloning). Average insert size 1.5 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 019. Note: this is a NIH MGC Library."

BASE COUNT	149 a	348 c	337 g	176 t	ORIGIN
Query Match	10.3%	Score 127;	DB 12;	Length 1010;	
Best Local Similarity	55.2;	Pred. No. 2.1e-16;			
Matches 294;	Conservative 0;	Mismatches 230;	Indels 9;	Gaps 2;	
Qy	27	CGGCCCCGAGGGGCGCGGGAGCGCGTGGCCCGCGCTGCGGCTTGGCGACGAGCGCG 86			
Db	82	CAGCAGCCCGGGCCCCCGGGCCAGCTCCAACCCGGGGCTGAGCTGTGACAGCGCCCGCT 141			
Qy	87	CTGCTCGCCCTTTCCCTTGGGGGCGCTGTGTCCGGTGACCGCTGTGTCCCTGTGCTCTGTT 146			
Db	142	GGGCGTGGACACTCGCTCTTGGGCCAAGGTGCTGTTACCGCGCTCTACGCACTCATCTG 201			
Qy	147	CGTCTCGGGGTGAGCGGCACGTG-----GTACCGTGATGCTGATCGGGCGCTACCG 200			
Db	202	GGCGCTGGGCGCGCGGGCAATGCGTGTGCGTGCACTGGTGTGAAGCGCGGGCGCG 261			
Qy	201	GGACATCGGACACCACTTGTATCTTGGGGAGCATGCGCCGTGTCCGACCTTACTCAT 260			
Db	262	GGCGCGGGGCGCTGCGCCACACGCTGCTCAGCTTGGCGCTCGCGGCGCTGCTGCTGCT 321			
Qy	261	CTGCTCGGGCTGCGGTTTGGACTGTACCGCTCTGGCGCTTCGGG---CCCTGGGTGTT 317			
Db	322	GCTGGTGGGCGTGCCTGGAGCTCTACAGCTTCTGTGTGTTCCACTACCCCTGGGTCTT 381			
Qy	318	CGGGCCGCTGTCTGTGCGCGCTTCCCTCTACGTGGGCGAGGGGTGCACCTACGCCACGCT 377			
Db	382	CGGCGACTTGGGCTTGC CGGGCTACTTCTGTCGACGAGTGTGCGCTTACGCCACGGT 441			
Qy	378	GCTGCA CATGACCGGGTCTACGGTTCGAGCGCTACCTGGCCATCTGCCCGCGCTCCGGC 437			
Db	442	GCTGAGCGTGGCAGCGCTGAGCGCGAGCGTGCCTAGCCGTGTGTCCAGAGCCCTTCGCTGC 501			
Qy	438	CCGGCTCTTGGTCAACCGGGCGCGGCTCCGCGGCTCATCGCTGTCTCTGGGGCGGTGGC 497			
Db	502	CCGCGAGCTGCTGACGCAACCCCGACCCGGTGTGTTGGCGCTCTCTGTGGGCGGCTC 561			
Qy	498	GCTGCTCTCTCGCGGTCCCTTCTTGTCTCTGGTGGGCGTGCAGCAGACCCCG 550			
Db	562	GCTCGGCTGCGCTCGCCATGCGCGCTCATCATGGGCGAAGCAGCAGCACTCG 614			

RESULT 9
 BG714306
 LOCUS
 DEFINITION
 634 bp mRNA linear EST 08-MAY-2001
 602669812F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4792730 5',
 mRNA sequence.
 ACCESSION
 BG714306
 VERSION
 BG714306.1 GI:13993237
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 634)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL
 Unpublished
 COMMENT
 Contact: Robert Strausberg, Ph.D.

found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLNL10671 row: 9 column: 03
 High quality sequence stop: 633.

```

FEATURES
source
Location/Qualifiers
1. .634
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4792730"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/clone_lib="NH_MGC_96"
/note="Organ: brain; Vector: pBluescript
pBluescript KS+; Site 1: BamHI; Site 2
); Oligo-dT primed using primer 5'-TTT
size-selected for average insert size
normalized to 10^5. This is a primary
for full-length clones and constructed
Cap-trapper method (Carninci, in prepar
constructed by M. Brownstein (NIMH/NGRR
Institutes of Health). Note: this is a
74 a 221 c 243 g 96 t
BASE COUNT

```

```

BASE COUNT      74 a      221 c      243 g      96 t
ORIGIN
Query Match      10.1%; Score 125.4; DB 10; Length 634;
Best Local Similarity 55.0%; Pred. No. 4.1e-16;
Matches 293; Conservative 0; Mismatches 231; Indels 9; Gaps 2;

Qy 27 CGGCCCCGAGGGGGCGCGGAGCGCGCGCTGCGCGCGCTGCGCGCTTGACGAGCGCG 86
    |||
Db 49 CAGCAGCCCGCGCCCCCGCGGCCAGCTCCAAACCCGGGGCTGAGCTGCGACCCCGGCT 108
    |||
Qy 87 CTGCTCGCCCTTTCCTCTGGGGCGCTGTGTCCCGGTGACCGCTGTGTGCTGTGCTGTT 146
    |||
Db 109 GGGCGTGGACACTCGCTCTGGGCCAAGTGTCTTTACCGCGCTCTACGCACTCATCTG 168
    |||
Qy 147 CGTGTCTGGGGTGAGCGGCAACGTG-----GTGACCGTGAATCTGATCGGGCGCTACCG 200
    |||
Db 169 GGCCTCTGGGCGCGCGGGCAATGCGCTGTGCGTGCACGTGTGTCTGAAGCGCGGGCCGG 228
    |||
Qy 201 GGACATGCGGACCACACCAACATGTATCTTGGGCAGCATGGCCGCTGTCCGACCTACTCAT 260
    |||
Db 229 GCGAGCGGGGCGCTGCGGCCACACACGTGCTCAGCTGCGCGCTCGCGGGCTGTGCTGCT 288
    |||
Qy 261 CTTGCTCGGGCTGCGGTTTGAACTGTACCGCTCTCGCGCTCGCGG-----CCCTGGGTGTT 317
    |||
Db 289 GCTGCTCGGCGTGCCTGGAGCTCTACAGCTCTGCTGTGGTTTCACCTACCCCTGGGCTCT 348
    |||
Qy 318 CGGGCGCGCTGCTCTGCGCGCTGTCCCTCTACGTGGGCGAGGGGTGCACCTACGCCACGCT 377
    |||
Db 349 CGGCGACCTGGGCTGCGCGGCTACTTACTTGTGTCACGAGCTGTGCGGCTACGCCAGGT 408
    |||
Qy 378 GCTTGCAATGACCGCGCTGAGGCTGAGCGCTACCTGGGCCATCTGCGGCGCGCTCGCGCG 437
    |||
Db 409 GCTGAGCGTGGCAGGCTTGAGCGCCGAGCGCTGCTAGCCGTGTGTCAGCGCCCTGCGTGC 468
    |||
Qy 438 CCGCGCTTTGGTCAACCGGGCGCGCGTTCGCGCGGCTCATCGCTGTGCTTGGGCGCGTGGC 497
    |||
Db 469 CCGCAGCTGCTGACGCGCAGCGCGGACCCGGTGGCTGTGGCGCTCTCTGTGGGCGCCCTC 528
    |||
Qy 498 GCTGCTCTTCGCGGTCCTCTTTGTTCTGTGTGGGCGTGCAGACGAGACCCCG 550
    |||
Db 529 GCTCGGCTCGCATGCCATGGCCGCTGATCATGTGGGCGAGAACGACGAACTCG 581
    |||

```

RESULT 10
BM808829
LOCUS
DEFINITION
 BM808829 1055 bp mRNA linear EST 05-MAR-2002
 AGENCOURT_6617918 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5734336
 5', mRNA sequence.
ACCESSION
 BM808829
VERSION
 BM808829.1 GI:19125652


```
Qy 501 GCTCTGCGCGTCCCTTCTTGTTCTTGTTGGGGTGGAGCAGGACCCCGGCACTCCGT 560
|||
Db 127 GCTGCGCGTGCACCAACACAGCGTGCACGGCATCGAGGSGTGGCGCCCAAGTTCGGC 68
|||
Qy 561 AGTCCCGGGCTCAATGGCACCAGCGGATCGCTCTCCGCTCTCGCTCGCTGCG 615
|||
Db 67 CGGAGTTTCTCTCGCCCATCTGCTGTAGGTCCCTCGAACCTCTGGTCTCTCG 13
|||

RESULT 15
BM808823 1603 bp mRNA linear EST 05-MAR-2002
LOCUS AGENCOURT 6617870 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5734333
5', mRNA sequence.
ACCESSION BM808823
VERSION BM808823.1 GI:19125646
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1603)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Invitrogen
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM12739 row: h column: 14
High quality sequence start: 84
High quality sequence stop: 514.
Location/Qualifiers
1. 1603
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5734333"
/tissue_type="hippocampus"
/lab_host="DH10B"
/clone_lib="NIH_MGC_124"
/notes="Organ: Brain; Vector: pCMV-SPORT6; Site_1: EcorV
(destroyed); Site 2: NotI; RNA source male hippocampus,
age 27. Library is oligo-dT primed and directionally
cloned (EcorV site is destroyed upon cloning). Average
insert size 1.4 kb, insert size range 0.9-4 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 012."
BASE COUNT 317 a 579 c 433 g 266 t 8 others
ORIGIN

Query Match 9.6%; Score 119; DB 12; Length 1603;
Best Local Similarity 54.5%; Pred. No. 1.1e-14;
Matches 286; Conservative 0; Mismatches 230; Indels 9; Gaps 2;

Qy 27 CGGCCCCGAGGGCGCGGAGCGCGCTGCGCCGCGCTTGCAGCAGCGCGC 86
|||
Db 131 CAGCAGCCGCGCGCGCGCGCGCTCAACCCGCGGCTGAGCTGAGCGCCGCT 190
|||
Qy 87 CTGCTCGCCCTTCCCTCGGGGCGCTGCTGCGCGCTGACCGCTGTGCTGTGCTGTT 146
|||
Db 191 GGGCGTGGACACTCGCTCTGGGCAAGGTGCTGTTACCGCGCTCTAGGCACATCATG 250
|||
Qy 147 CGTCTCGGGGTGAGCGGCAACGTG-----GTACCGTGTATGTCGGGCGCTACCG 200
|||
Db 251 GGCCTGGGCGCGGCGCAATGCGCTGTCCGTGCACTGCTGTGTAAGCGCGCGCGG 310
|||
```

```
Qy 201 GGACATGCGGACCAACCAACTTGTACTTGGGCGAGCATGSCGCTGTCCGACCTACTCAT 260
|||
Db 311 GCGGGGCGGCGCGCTGCGGCACACAGTGTCTAGCCTGGCGCTCGCGGCGCTGTCTGCT 370
|||
Qy 261 CTGCTCGGGCTGCGGTTGACCTGTACCGCTC---TGGCGGTGCGGGCGCTGGGGTGT 317
|||
Db 371 GCTGGTGGGCTGCGGTTGGAGCTCTACAGCTTCGTGTGGTTCCACTTACCCCTGGGTCTT 430
|||
Qy 318 CGGGCGCTGCTCTGCGCGCTGTCCCTTACGTGGGCGAGGGTGCACCTACGCCACGCT 377
|||
Db 431 CGGGACCTGGGCTGCGCGGCTACTACTTCTGTGCAGAGCTGTGCGCTACGCCACGGT 490
|||
Qy 378 GCTGCACATGACCGCGCTCAGCGTCGAGCGCTACCTGGCCATCTGCGCGCGCTTCCGCGC 437
|||
Db 491 GCTGAGCGTGGCAGGCTGAGCGCCGAGCGCTGCTAGCCGTGTGCCAGCCCTGCGTGC 550
|||
Qy 438 CCGGCTTGGTCAACCGGCGCGCGCTCCGCGCGCTCATCGCTGTGCTCTGGGCGCTGC 497
|||
Db 551 CCGCAGCTGCTGACGCGCACCGCGGACCCCGTGGTGGCGCTCTCTGTTGGGCGGCTC 610
|||
Qy 498 GCTGCTCTGCGCGTCCCTTCTTGTTCCTGGTGGGCGTCGAGCA 542
|||
Db 611 GCTCGGCTCCACCTGCCCCATGCGCCAGACATCATGGGCGCAGAAGCA 655
|||
```

Search completed: January 1, 2004, 02:52:40
Job time : 2913.48 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 1, 2004, 06:20:00 ; Search time 52.1454 Seconds
(without alignments)
1254.097 Million cell updates/sec

Title: US-09-719-485-3

Perfect score: 2155

Sequence: 1 MGSPWNSGDEGAREPPWP.....DTGGDTVGYTETSANVKTWG 412

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq_19Jun03:*

```

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description
1	2155	100.0	412	21 AAY54145
2	2155	100.0	412	22 AAB62652
3	2155	100.0	412	22 AAB68476
4	2155	100.0	412	23 AAB68478
5	2155	100.0	412	23 AAB68479
6	2155	100.0	412	24 AAB68479
7	2149	99.7	412	21 AAB62652
8	2141	99.4	412	23 AAB62652
9	1747	81.1	400	22 AAB68477

10	1581	73.4	386	21	AAY54146	Amino acid sequenc
11	1581	73.4	386	22	AAB62653	Short form of motil
12	1149.5	53.3	271	22	AAB68476	Amino acid sequenc
13	907	42.1	363	21	AAY54147	The puffer fish mo
14	907	42.1	363	22	AAB68479	Amino acid sequenc
15	865.5	40.2	349	21	AAY69293	A canine growth ho
16	863.5	40.1	366	21	AAY90832	Human G protein-co
17	863.5	40.1	366	21	AAY70345	Human G protein-co
18	863.5	40.1	366	22	AAB97376	Rat growth hormone
19	863.5	40.1	366	22	AAB62650	Human G-protein co
20	863.5	40.1	366	23	ABB09534	Human ghrelin rece
21	863	40.0	353	18	AAW19608	Pig growth hormone
22	863	40.0	353	18	AAW19215	Swine growth hormo
23	861.5	40.0	364	21	AAY54565	A mouse growth hor
24	861.5	40.0	364	22	AAB97377	Rat growth hormone
25	860	39.9	361	18	AAW19217	Human growth hormo
26	860	39.9	362	18	AAW19610	Human growth hormo
27	857.5	39.8	364	18	AAW19613	Rat growth hormone
28	857.5	39.8	366	21	AAY90866	Human mutant G pro
29	845.5	39.2	364	18	AAW19220	Rat growth hormone
30	674.5	31.3	271	18	AAW19612	Human growth hormo
31	674.5	31.3	271	18	AAW19219	Human growth hormo
32	644.5	29.9	289	18	AAW19611	Human growth hormo
33	644.5	29.9	289	24	ABP81828	Human growth hormo
34	641	29.7	289	18	AAW19609	Pig growth hormone
35	639.5	29.7	289	18	AAW19216	Swine growth hormo
36	639.5	29.7	289	18	AAW19218	Human growth hormo
37	501	23.2	418	17	ABP8562	Human neurotensin
38	497	23.1	418	24	ABP81861	Human neurotensin
39	497	23.1	542	22	ABG12316	Novel human diagno
40	491	22.8	418	22	ABP56371	Non-endogenous hum
41	486.5	22.6	445	22	AAU68523	Human novel cytol
42	486.5	22.6	445	22	AAU68566	Human novel cytol
43	485.5	22.5	403	21	AAY90638	Human G protein-co
44	485.5	22.5	403	21	AAV44642	Human growth hormo
45	485.5	22.5	403	22	AAB99185	Human PK-3. Homo

ALIGNMENTS

RESULT 1

AAV54145

ID AAY54145 standard; Protein; 412 AA.

XX AC AAY54145;

XX DT 27-MAR-2000 (first entry)

XX DE Amino acid sequence of the motilin receptor splice variant MTL-R1A.

XX KW Motilin receptor; G-protein coupled receptor; MTL-R1; GPR38;
 KW Spliced form; MTL-R1A; MTL-R1B; gastric motility disorder;
 KW functional defect; neurological disorder; scleroderma; colonoscopy;
 KW paraneoplastic syndrome; radiation induced dysmotility; diabetes;
 KW infection; stress-related motility disorder; psychogenic disorder;
 KW gastroparesis; gastro-oesophageal reflux disease; constipation;
 KW chronic idiopathic pseudo obstruction; acute faecal impaction;
 KW postoperative ileus; gallstones; infantile colic; diarrhoea;
 KW irritable bowel syndrome; non-ulcer dyspepsia; non-cardiac chest pain;
 KW endoscopy; duodenal intubation.

XX OS Homo sapiens.

XX PN WO9964436-A1.

XX PD 16-DEC-1999.

XX PF 08-JUN-1999; 99WO-US12773.

XX PR 12-JUN-1998; 98US-0089098.

XX PA (MERI) MERCK & CO INC.

XX Feighner SD, Patchett AA, Tan C, McKee K, MacNeil D, Howard AD;
 PI Pong S, Smith RG;
 XX
 XX
 DR WPI; 2000-105868/09.
 DR N-PSDB; AAZ45403.
 XX
 PT Novel receptor protein for screening compounds used in treating
 PT irritable bowel syndrome, constipation and other gastric conditions -
 XX
 PS Claim 3; Fig 3; 44pp; English.
 XX
 CC The present sequence represents splice variant MTL-R1A of the motilin
 CC receptor. The gene encodes a G-protein coupled receptor, and is
 CC designated MTL-R1 (also GPR38). Two spliced forms of GPR38 exist,
 CC MTL-R1A (see AAY54145) and MTL-R1B (see AAY54146). MTL-R1A is a
 CC functional seven transmembrane domain form, and MTL-R1B is a truncated
 CC five transmembrane domain. The MTL-R1 proteins are used to identify
 CC agonists and antagonists which can be used for treating gastric motility
 CC disorders, functional defects, disorders secondary to neurological
 CC disorders e.g. scleroderma, paraneoplastic syndromes radiation induced
 CC dysmotility, diabetes, infections, stress-related motility disorders,
 CC psychogenic disorders, gastroparesis, gastro-oesophageal reflux disease,
 CC constipation, chronic idiopathic pseudo obstruction, acute faecal
 CC impaction, postoperative ileus, gallstones, infantile colic, irritable
 CC bowel syndrome, non-ulcer dyspepsia, non-cardiac chest pain and
 CC diarrhoea. They can also be used in the preparation for colonoscopy,
 CC endoscopy and duodenal intubation.
 XX
 SQ Sequence 412 AA;
 Query Match 100.0%; Score 2155; DB 21; Length 412;
 Best Local Similarity 100.0%; Pred. No. 1.2e-194;
 Matches 412; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGSPWNGSDGPEGAREPPWALPPCDERRCSPPLGALVPVTAIVCLFVVGSGNVTV 60
 DB 1 MGSPWNGSDGPEGAREPPWALPPCDERRCSPPLGALVPVTAIVCLFVVGSGNVTV 60
 QY 61 MLIGRYDMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPFGLFLVGVGEG 120
 DB 61 MLIGRYDMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPFGLFLVGVGEG 120
 QY 121 CTYATLLHMTALSVERYLAICRPLARVLVTRRRVRLAIVLWAVALLSAGPFLVGVGVE 180
 DB 121 CTYATLLHMTALSVERYLAICRPLARVLVTRRRVRLAIVLWAVALLSAGPFLVGVGVE 180
 QY 181 QDPGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAEAAALFSRECRSPA 240
 DB 181 QDPGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAEAAALFSRECRSPA 240
 QY 241 QLGNALVMLWTTAYFFLPFLCLISILYGLIGRELWSSRPRLGPAASGRGRGHQTVRL 300
 DB 241 QLGNALVMLWTTAYFFLPFLCLISILYGLIGRELWSSRPRLGPAASGRGRGHQTVRL 300
 QY 301 LVVLAFLICWLPFHVGRIIYINTEDSRMVYFSQYFNIVALQLFYLSASINPLYNLSK 360
 DB 301 LVVLAFLICWLPFHVGRIIYINTEDSRMVYFSQYFNIVALQLFYLSASINPLYNLSK 360
 QY 361 KYRAAFLKLLARKSRPRGHRSDTAGEVAGDTGGDTGVGTSTANVKTWG 412
 DB 361 KYRAAFLKLLARKSRPRGHRSDTAGEVAGDTGGDTGVGTSTANVKTWG 412
 RESULT 2
 AAB62652
 ID AAB62652 standard; Protein; 412 AA.
 XX
 AC AAB62652;
 XX
 DT 23-JUL-2001 (first entry)
 XX
 DE Long form of motilin receptor, GPR-38A isoform.

XX zsig33; signal transduction; hormone; enzyme; neural development;
 KW gastric contractility; nutrient uptake; digestive; pancreatic; human;
 KW insulin-like growth factor-I; growth hormone; bone; gastrointestinal;
 KW glucose; osteopathic; anorectic; vulnerable; immunomodulatory; GHS-R;
 KW G-protein coupled receptor; motilin receptor; GPR38; GPR-38A; isoform.
 XX
 OS Homo sapiens.
 XX
 XX WO200138355-A2.
 XX
 XX 31-MAY-2001.
 XX
 XX 22-NOV-2000; 2000WO-US32074.
 XX
 XX 22-NOV-1999; 99US-0166765.
 PR
 XX (ZYMO) ZYMOGENETICS INC.
 XX
 XX Sheppard PO, Jaspers SR, Deisher TA, Bishop PD;
 DR WPI; 2001-355879/37.
 DR N-PSDB; AAF83683.
 XX
 PT Forming reversible peptide receptor complex for purifying cell and
 PT peptides, stimulating signal transduction and modulating hormone
 PT secretion, involves contacting a receptor with zsig33 polypeptide -
 XX
 XX Disclosure; Page 102-104; 11pp; English.
 XX
 CC The invention relates to a method of forming a reversible peptide-
 CC receptor complex that involves providing an immobilized receptor, and
 CC contacting the receptor with a zsig33 peptide (comprising residues 24-37
 CC of AAB6249), where the receptor binds to the zsig33 peptide. The method
 CC is useful for purifying cells, purifying a peptide, stimulating signal
 CC transduction in a cell expressing a receptor. It is also useful for
 CC modulating secretion of hormones, neural development and/or utilization,
 CC gastric contractility, nutrient uptake, secretion of digestive and
 CC pancreatic enzymes and hormones, secretion of insulin-like growth factor
 CC -I, secretion of non-zsig33 proteins. It is useful for modulating growth
 CC hormone secretion in a mammal having a disease associated with abnormal
 CC levels of growth hormone, such as osteoporosis, bone repair, bone
 CC remodeling, low osteoblast levels, cartilage repair and remodeling,
 CC skeletal dysplasia, immune suppression, obesity, growth retardation,
 CC protein catabolic responses after surgery, cachexia, protein loss,
 CC dwarfism, wound healing and ovulation induction, treating a mammal having
 CC a metabolic disorder requiring neurological feedback, such as satiety
 CC regulation, glucose absorption and metabolism and neuropathy-associated
 CC gastrointestinal disorders, and stimulating glucose-induced insulin
 CC release in a mammal. The present sequence represents the long form of
 CC motilin receptor, GPR-38A (one of the two isoforms of GPR38 which result
 CC from alternative splicing). GPR38 has homology to the human G-protein
 CC coupled receptor, GHS-R.
 XX
 SQ Sequence 412 AA;
 Query Match 100.0%; Score 2155; DB 22; Length 412;
 Best Local Similarity 100.0%; Pred. No. 1.2e-194;
 Matches 412; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGSPWNGSDGPEGAREPPWALPPCDERRCSPPLGALVPVTAIVCLFVVGSGNVTV 60
 DB 1 MGSPWNGSDGPEGAREPPWALPPCDERRCSPPLGALVPVTAIVCLFVVGSGNVTV 60
 QY 61 MLIGRYDMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPFGLFLVGVGEG 120
 DB 61 MLIGRYDMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPFGLFLVGVGEG 120
 QY 121 CTYATLLHMTALSVERYLAICRPLARVLVTRRRVRLAIVLWAVALLSAGPFLVGVGVE 180
 DB 121 CTYATLLHMTALSVERYLAICRPLARVLVTRRRVRLAIVLWAVALLSAGPFLVGVGVE 180
 QY 181 QDPGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAEAAALFSRECRSPA 240

```

|||||
181 QDPGISVVPGLNGTARIASSPLASSPPLMLSRAPPSPGPGPETAEEAALFSRECRPSA 240
|||||
241 QLGALRVMLVMTTAYTFPLPCLLSILYGLIGRELWSSRRPLRGPAAASGRGRHRTQVRVL 300
|||||
241 QLGALRVMLVMTTAYTFPLPCLLSILYGLIGRELWSSRRPLRGPAAASGRGRHRTQVRVL 300
|||||
301 LVVVLAFIICWLPFHVGRIIYINTEDSRMMYFSQYFNIVALQLFYLSASINPILYNLISK 360
|||||
301 LVVVLAFIICWLPFHVGRIIYINTEDSRMMYFSQYFNIVALQLFYLSASINPILYNLISK 360
|||||
361 KYRAAFAKLLARKSRPRGHRSDRTAGEVAGDTGGDTVGYTETSANVKTWG 412
|||||
361 KYRAAFAKLLARKSRPRGHRSDRTAGEVAGDTGGDTVGYTETSANVKTWG 412

RESULT 3
AAB68478
ID AAB68478 standard; Protein; 412 AA.
XX
AC AAB68478;
XX
DT 23-JUL-2001 (first entry)
XX
DE Amino acid sequence of a human motilin receptor polypeptide.
XX
KW Motilin receptor; gastrointestinal disease; gastric motility disorder;
KW gastroparesis; irritable bowel syndrome; diarrhoea.
XX
OS Homo sapiens.
XX
XX WO200132710-A1.
XX
XX 10-MAY-2001.
XX
XX 25-OCT-2000; 2000WO-US29426.
XX
XX 29-OCT-1999; 99US-01622264.
XX
XX (MERI ) MERCK & CO INC.
XX
XX Tan C, McKee K;
XX
XX WPI; 2001-343479/36.
XX
XX N-PSDB; AAF85449.
XX
PT Novel polypeptides related to dog and rabbit motilin receptor
PT polypeptide, comprising unique regions from dog and motilin receptor
PT amino acid sequence, useful for identifying compounds for treating
PT diarrhoea in humans
XX
XX Disclosure; Page 32-33; 42pp; English.
XX
CC The present sequence represents a human motilin receptor polypeptide.
CC The specification describes a unique sequence present in exon 1 of
CC the dog motilin receptor, which is not present in human or Sphaeroides
CC nephelus 75E7 motilin receptor sequences. The unique nucleic acid
CC sequence is useful for measuring the ability of a compound to affect
CC motilin receptor activity. Motilin receptor polynucleotides and
CC polypeptides are used to identify therapeutic compounds which are
CC useful for treating gastrointestinal diseases and disorders such as
CC gastric motility disorders, gastroparesis, irritable bowel syndrome,
CC and diarrhoea.
XX
XX Sequence 412 AA;
XX
Query Match 100.0%; Score 2155; DB 22; Length 412;
Best Local Similarity 100.0%; Pred. No. 1.2e-194;
Matches 412; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGSPWNGSDGPEGAREPPWALPCDERRCSPFPLGALVPVTAVCLCLFVVGSGNVVTV 60
Db 1 MGSPWNGSDGPEGAREPPWALPCDERRCSPFPLGALVPVTAVCLCLFVVGSGNVVTV 60

```

```

QY 61 MLIGRYRDMRTTNTNLYLGSMAVSDLLILGLPPDLVRLWRSRPMWFGPLLCRLSLYVGEG 120
|||||
61 MLIGRYRDMRTTNTNLYLGSMAVSDLLILGLPPDLVRLWRSRPMWFGPLLCRLSLYVGEG 120
|||||
121 CTYATLLHMTALSVRYLAICRPLRARVLTTRRVRALIAVLWAVALLSAGPPLFLVGVE 180
|||||
121 CTYATLLHMTALSVRYLAICRPLRARVLTTRRVRALIAVLWAVALLSAGPPLFLVGVE 180
|||||
181 QDPGISVVPGLNGTARIASSPLASSPPLMLSRAPPSPGPGPETAEEAALFSRECRPSA 240
|||||
181 QDPGISVVPGLNGTARIASSPLASSPPLMLSRAPPSPGPGPETAEEAALFSRECRPSA 240
|||||
241 QLGALRVMLVMTTAYTFPLPCLLSILYGLIGRELWSSRRPLRGPAAASGRGRHRTQVRVL 300
|||||
241 QLGALRVMLVMTTAYTFPLPCLLSILYGLIGRELWSSRRPLRGPAAASGRGRHRTQVRVL 300
|||||
301 LVVVLAFIICWLPFHVGRIIYINTEDSRMMYFSQYFNIVALQLFYLSASINPILYNLISK 360
|||||
301 LVVVLAFIICWLPFHVGRIIYINTEDSRMMYFSQYFNIVALQLFYLSASINPILYNLISK 360
|||||
361 KYRAAFAKLLARKSRPRGHRSDRTAGEVAGDTGGDTVGYTETSANVKTWG 412
|||||
361 KYRAAFAKLLARKSRPRGHRSDRTAGEVAGDTGGDTVGYTETSANVKTWG 412

RESULT 4
ABG30936
ID ABG30936 standard; Protein; 412 AA.
XX
AC ABG30936;
XX
DT 21-OCT-2002 (first entry)
XX
DE Human G protein-coupled receptor 38 (GPR38).
XX
KW Human; G protein-coupled receptor 38; receptor; GPR38;
KW Alzheimer's disease; Parkinson's disease; ulcerative colitis;
KW Crohn's disease; Hodgkin's disease; glioblastoma; breast carcinoma;
KW colon carcinoma; lung small cell carcinoma; lung adenocarcinoma;
KW pancreatic small cell carcinoma; pancreatic adenocarcinoma.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX 15..31
XX Region /note= "Antigenic fragment"
XX Region 217..232
XX /note= "Antigenic fragment"
XX Region 276..291
XX /note= "Antigenic fragment"
XX Region 373..388
XX /note= "Antigenic fragment"
XX WO200257791-A2.
XX
XX 25-JUL-2002.
XX
XX 29-NOV-2001; 2001WO-US45219.
XX
XX 29-NOV-2000; 2000US-250251P.
XX
XX 30-NOV-2000; 2000US-250452P.
XX
XX (LIFE-) LIFESPAN BIOSCIENCES INC.
XX
XX Brown JP, Burmer GC, Roush CL, Kulander BG;
XX WPI; 2002-566812/60.
XX
XX N-PSDB; ABK90132.
XX
XX Assay for detecting Alzheimer's disease, Parkinson's disease,
XX ulcerative colitis, Crohn's disease, Hodgkin's disease, glioblastoma or
XX carcinoma, comprises using a binding partner for G protein coupled

```

PT receptor 38 -
 PS Disclosure; Fig 2; 112pp; English.
 XX
 CC The present invention relates to a new assay method that involves
 CC contacting a binding partner specific for G protein coupled receptor
 CC (GPR) 38 with specific cells. The method of the invention is useful for
 CC the detection of an increased risk of Alzheimer's disease, Parkinson's
 CC disease, ulcerative colitis, Crohn's disease, Hodgkin's disease,
 CC glioblastoma, or carcinoma. GPR 38 is used to manufacture a medicament
 CC for inhibiting, treating or preventing Alzheimer's disease, Parkinson's
 CC disease, ulcerative colitis, Crohn's disease, Hodgkin's disease,
 CC glioblastoma, breast carcinoma, colon carcinoma, lung small cell
 CC carcinoma, lung adenocarcinoma, pancreatic small cell carcinoma, and
 CC pancreatic adenocarcinoma. An agonist or antagonist to GPR 38 are used
 CC to manufacture a medicament able to reduce the symptoms of these
 CC diseases. Nucleic acids encoding GPR 38 can also be used to treat the
 CC diseases. The present amino acid sequence represents the human G
 CC protein-coupled receptor 38 (GPR38) of the invention.
 XX
 SQ Sequence 412 AA;
 Query Match 100.0%; Score 2155; DB 23; Length 412;
 Best Local Similarity 100.0%; Pred. No. 1.2e-194;
 Matches 412; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGSPWNGSDGPEGAREPPWALPPCDERRCSPPLGALVPTAVCLCLFVVGSGNVTV 60
 DB 1 MGSPWNGSDGPEGAREPPWALPPCDERRCSPPLGALVPTAVCLCLFVVGSGNVTV 60
 QY 61 MLIGRYDMRTTNNLYLGSMVSDLLILGLPFDLYELWRSRPPWVFGPFLCRSLYVGE 120
 DB 61 MLIGRYDMRTTNNLYLGSMVSDLLILGLPFDLYELWRSRPPWVFGPFLCRSLYVGE 120
 QY 121 CTYATLLHMTALSVRYLAICRPLARVLTTRRRVRLIAVLWAVALLSAGPFLVGV 180
 DB 121 CTYATLLHMTALSVRYLAICRPLARVLTTRRRVRLIAVLWAVALLSAGPFLVGV 180
 QY 181 QDPGISVVGNGTARTASSPLASSPPLWLSRPPSPGPTAEAAALFSRRCRPSA 240
 DB 181 QDPGISVVGNGTARTASSPLASSPPLWLSRPPSPGPTAEAAALFSRRCRPSA 240
 QY 241 QLGLARVLMVWTAYFFLPCLCSILYGLGRLWSSRRPLRGAASGRGRHQTVRVL 300
 DB 241 QLGLARVLMVWTAYFFLPCLCSILYGLGRLWSSRRPLRGAASGRGRHQTVRVL 300
 QY 301 LVVVLAPIICWLPFHVGRIIYINTDSRMVYFQYFNIVALQIFYLSASINPILYNLISK 360
 DB 301 LVVVLAPIICWLPFHVGRIIYINTDSRMVYFQYFNIVALQIFYLSASINPILYNLISK 360
 QY 361 KYRAAFAKLLARKSRPRGHRSDTAGVAGTGGDTGVTGTSANVKTGM 412
 DB 361 KYRAAFAKLLARKSRPRGHRSDTAGVAGTGGDTGVTGTSANVKTGM 412

RESULT 5
 ABP81993
 ID ABP81993 standard; Protein; 412 AA.
 XX
 AC ABP81993;
 XX
 DT 04-MAR-2003 (first entry)
 XX
 DE Human motilin receptor GPR38 protein SEQ ID NO:473.
 XX
 KW G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
 KW G protein-coupled receptor modulator; antibody; immune-related disease;
 KW growth-related disease; cell regeneration-related disease; AIDS; cancer;
 KW immunological-related cell proliferative disease; autoimmune disease;
 KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
 KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
 KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
 KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;

KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
 KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
 KW ulcer.
 XX Homo sapiens.
 OS
 XX WO200261087-A2.
 PN
 XX WO200261087-A2.
 PD
 XX 08-AUG-2002.
 XX
 PF 19-DEC-2001; 2001WO-US50107.
 XX
 XX 19-DEC-2000; 2000US-257144P.
 PR
 XX (LIFE-) LIFESPAN BIOSCIENCES INC.
 PA
 XX Burmer GC, Roush CL, Brown JP;
 PI
 XX WPI; 2003-046718/04.
 DR
 XX N-PSDB; ABZ42842.
 XX
 XX New isolated antigenic peptides e.g., for G protein-coupled receptors
 PT (GPCR); useful for diagnosing and designing drugs for treating
 PT conditions in which GPCRs are involved, e.g. AIDS, Alzheimer's disease,
 PT cancer or autoimmune diseases -
 XX
 PS Disclosure; Fig 1; 523pp; English.
 XX
 CC The present invention describes antigenic peptides (I) comprising:
 CC (a) any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
 CC acids. Also described: (1) an assay for the detection of a particular
 CC G protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
 CC and (2) an isolated antibody having high specificity and high affinity
 CC or avidity for a particular GPCR. (I) can be used as GPCR modulators and
 CC in gene therapy. The antigenic peptides for GPCRs are useful in detecting
 CC an antibody against a particular GPCR, and in the production of specific
 CC antibodies. The peptides and antibodies are also useful for detecting the
 CC presence or absence of corresponding GPCRs. The antigenic peptides for
 CC GPCRs and antibodies are useful for diagnosing and designing drugs for
 CC treating immune-related diseases, growth-related diseases, cell
 CC regeneration-related disease, immunological-related cell proliferative
 CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
 CC atherosclerosis, bacterial, fungal, protozoan or viral infections,
 CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
 CC inflammation, allergies, Crohn's disease, diabetes, graft versus host
 CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
 CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
 CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
 CC hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
 CC any other disorder in which GPCRs are involved. The antibodies may be
 CC used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode
 CC GPCR proteins given in ABP81675 to ABP82018, which are used in the
 CC exemplification of the present invention.
 XX
 SQ Sequence 412 AA;
 Query Match 100.0%; Score 2155; DB 24; Length 412;
 Best Local Similarity 100.0%; Pred. No. 1.2e-194;
 Matches 412; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGSPWNGSDGPEGAREPPWALPPCDERRCSPPLGALVPTAVCLCLFVVGSGNVTV 60
 DB 1 MGSPWNGSDGPEGAREPPWALPPCDERRCSPPLGALVPTAVCLCLFVVGSGNVTV 60
 QY 61 MLIGRYDMRTTNNLYLGSMVSDLLILGLPFDLYELWRSRPPWVFGPFLCRSLYVGE 120
 DB 61 MLIGRYDMRTTNNLYLGSMVSDLLILGLPFDLYELWRSRPPWVFGPFLCRSLYVGE 120
 QY 121 CTYATLLHMTALSVRYLAICRPLARVLTTRRRVRLIAVLWAVALLSAGPFLVGV 180
 DB 121 CTYATLLHMTALSVRYLAICRPLARVLTTRRRVRLIAVLWAVALLSAGPFLVGV 180
 QY 181 QDPGISVVGNGTARTASSPLASSPPLWLSRPPSPGPTAEAAALFSRRCRPSA 240

Db 181 QDPGIVVPLNGTARIASSPLASPPMLWSRRPSPGPTAEAAALFSRECRPSA 240
Qy 241 QLGALRVMLWTTTAYFFLPFLCLSLYLGLIGRELWSSRRPLRGPAASGRERGHRTQTVRVL 300
Db 241 QLGALRVMLWTTTAYFFLPFLCLSLYLGLIGRELWSSRRPLRGPAASGRERGHRTQTVRVL 300
Qy 301 LVVLAFLIICWLPFHVGRIIYINTEDSRMMYFSQYFNIVALQLPYLSASINPILYNLISK 360
Db 301 LVVLAFLIICWLPFHVGRIIYINTEDSRMMYFSQYFNIVALQLPYLSASINPILYNLISK 360
Qy 361 KYRAAFAKLLILARKSRPRGFRHSRDTAGEVAGDTGGDTVGVTETSANVKTMG 412
Db 361 KYRAAFAKLLILARKSRPRGFRHSRDTAGEVAGDTGGDTVGVTETSANVKTMG 412

RESULT 6

AAG65822 ID AAG65822 standard; Protein; 501 AA.
XX AC AAG65822;

DT 30-JAN-2002 (first entry)
XX DE Human GPR38 variant GPR38V polypeptide.

XX KW GPR38V; variant; antibacterial; cytostatic; analgesic; antiasthmatic;
XX KW anti-Parkinsonian; hypertensive; hypotensive; antidiabetic; osteopathic;
XX KW antiallergic; antimigraine; neuroleptic; nootropic; anticonvulsant;
XX KW antiulcer; antiemetic; cardiac; vaccine; human.

XX OS Homo sapiens.

XX PN WO200164836-A2.

XX PD 07-SEP-2001.

XX PF 28-FEB-2001; 2001WO-US06277.

XX PR 01-MAR-2000; 2000US-0516315.

XX PA (SMIK) SMITHKLINE BEECHAM CORP.

XX PI Elshourbagy N, Shabon U;

XX DR WPI: 2001-638956/73.

XX DR N-PSDB; AAI65989.

XX New human GPR38V polypeptide and polynucleotide, useful for treating
PT e.g. bacterial, fungal, protozoal and viral infections, cancers or
PT allergies, as vaccines, and for identifying agonists and antagonists
PT potentially useful in therapy -

XX FS Claim 1; Page 26; 32pp; English.

XX This represents a human GPR38 variant (GPR38V) polypeptide. GPR38V can
CC be expressed by standard recombinant methodology. The polynucleotides and
CC polypeptides are used in the treatment of bacterial, fungal, protozoal
CC and viral infections, infections caused by HIV-1 or HIV-2, pain, cancers,
CC diabetes, obesity, anorexia, asthma, Parkinson's disease, acute heart
CC failure, hypertension, urinary retentions, osteoporosis, allergies,
CC ulcers, migraine, psychotic and neurological disorders, or dyskinesias.
CC They are also useful for identifying agonists and antagonists that are
CC potentially useful in therapy, as vaccines to induce immunological
CC response in a mammal. The polypeptides may also be used as immunogens to
CC produce antibodies immunospecific for the polypeptides, and to identify
CC membrane bound or soluble receptors.

XX SQ Sequence 501 AA;

Query Match 100.0%; Score 2155; DB 22; Length 501;
Best Local Similarity 100.0%; Pred. No. 1.5e-194;
Matches 412; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGSPWNGSDGPEGARPPWPALPPCDERRCSPFLGALVPVTAVCCLFVVGSGNVVTV 60
Db 90 MGSPWNGSDGPEGARPPWPALPPCDERRCSPFLGALVPVTAVCCLFVVGSGNVVTV 149
Qy 61 MLIGRYDMRTTNNLYLGSMAVSDLLILGLPDLRLWRSRPWVPGPLICRLSLYVGE 120
Db 150 MLIGRYDMRTTNNLYLGSMAVSDLLILGLPDLRLWRSRPWVPGPLICRLSLYVGE 209
Qy 121 CTYATLLHMTALSVERYLAICRPLRARVLVTRRRVRALIAVLWAVALLSAGPPLFLVGV 180
Db 210 CTYATLLHMTALSVERYLAICRPLRARVLVTRRRVRALIAVLWAVALLSAGPPLFLVGV 269
Qy 181 QDPGISVVPGLNGTARIASSPLASPPMLWSRRPSPGPTAEAAALFSRECRPSA 240
Db 270 QDPGISVVPGLNGTARIASSPLASPPMLWSRRPSPGPTAEAAALFSRECRPSA 329
Qy 241 QLGALRVMLWTTTAYFFLPFLCLSLYLGLIGRELWSSRRPLRGPAASGRERGHRTQTVRVL 300
Db 330 QLGALRVMLWTTTAYFFLPFLCLSLYLGLIGRELWSSRRPLRGPAASGRERGHRTQTVRVL 389
Qy 301 LVVLAFLIICWLPFHVGRIIYINTEDSRMMYFSQYFNIVALQLPYLSASINPILYNLISK 360
Db 390 LVVLAFLIICWLPFHVGRIIYINTEDSRMMYFSQYFNIVALQLPYLSASINPILYNLISK 449
Qy 361 KYRAAFAKLLILARKSRPRGFRHSRDTAGEVAGDTGGDTVGVTETSANVKTMG 412
Db 450 KYRAAFAKLLILARKSRPRGFRHSRDTAGEVAGDTGGDTVGVTETSANVKTMG 501

RESULT 7
AAB02854 ID AAB02854 standard; Protein; 412 AA.
XX AC AAB02854;
XX DT 22-AUG-2000 (first entry)
XX DE Human G protein coupled receptor hGPR38 (V297K) protein SEQ ID NO:130.
XX KW Human; G protein coupled receptor; GPCR; transmembrane receptor;
XX KW identification; agonist; screening; therapeutic; pharmaceutical;
XX KW mutant.
XX OS Homo sapiens.
XX OS Synthetic.
XX PN WO200022131-A2.
XX PD 20-APR-2000.
XX PF 13-OCT-1999; 99WO-US24065.
XX PR 13-OCT-1998; 98US-0170496.
XX PR 12-NOV-1998; 98US-0108029.
XX PR 20-NOV-1998; 98US-0109213.
XX PR 27-NOV-1998; 98US-0110060.
XX PR 16-FEB-1999; 99US-0120416.
XX PR 26-FEB-1999; 99US-0121852.
XX PR 12-MAR-1999; 99US-0123944.
XX PR 12-MAR-1999; 99US-0123945.
XX PR 12-MAR-1999; 99US-0123946.
XX PR 12-MAR-1999; 99US-0123948.
XX PR 12-MAR-1999; 99US-0123949.
XX PR 12-MAR-1999; 99US-0123951.
XX PR 28-MAY-1999; 99US-0136436.
XX PR 28-MAY-1999; 99US-0136437.
XX PR 28-MAY-1999; 99US-0136439.
XX PR 28-MAY-1999; 99US-0137127.
XX PR 28-MAY-1999; 99US-0137131.
XX PR 30-JUN-1999; 99US-0137567.
XX PR 27-AUG-1999; 99US-0141448.
XX PR 27-AUG-1999; 99US-0151114.

PR 03-SEP-1999; 99US-0152524.
 PR 29-SEP-1999; 99US-0156633.
 PR 29-SEP-1999; 99US-0156555.
 PR 29-SEP-1999; 99US-0156634.
 XX
 PA (AREN-) ARENA PHARM INC.
 XX
 PI Behan DP, Lehmann-Bruinsma K, Chalmers DT, Chen R, Dang HT;
 PI Gore M, Liaw CW, Lin I, Lowitz K, White C;
 XX
 DR WPI; 2000-317986/27.
 DR N-PSDB; AAA46116.
 XX
 PT Non-endogenous, human G protein-coupled receptors for screening
 PT receptor, inverse or partial agonists useful as therapeutic agents -
 XX
 PS Example 2; Page 168-169; 187pp; English.
 XX
 CC The present invention describes transmembrane receptors, preferably
 CC human G protein coupled receptors (GPCR), for which the endogenous
 CC ligand is unknown (orphan GPCR receptors). More specifically the present
 CC invention relates to non-endogenous, constitutively activated versions
 CC of a human GPCR. These non-endogenous human GPCRs can be useful for
 CC the direct identification of candidate compounds as receptors agonists,
 CC inverse agonists or partial agonists for use as pharmaceutical agents.
 CC AAA46017 to AAA46126 and AAB02825 to AAB02859 represent sequences used in
 CC the exemplification of the present invention.
 XX
 SQ Sequence 412 AA;

Query Match 99.7%; Score 2149; DB 21; Length 412;
 Best Local Similarity 99.8%; Pred. No. 4.5e-194;
 Matches 411; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 MGSPWNGSDGPEGAREPPWPPALPPCDERRCSPPPLGALVPVTAACLVFVVGSGNVTV 60
 Db 1 MGSPWNGSDGPEGAREPPWPPALPPCDERRCSPPPLGALVPVTAACLVFVVGSGNVTV 60
 Qy 61 MLIGRYDRMTTNNLYLGSMAVSDLLILLGLPPDLYRLWRSRPPWFGPLCRLSLYVGE 120
 Db 61 MLIGRYDRMTTNNLYLGSMAVSDLLILLGLPPDLYRLWRSRPPWFGPLCRLSLYVGE 120
 Qy 121 CTYATLLHMTALSVERYLAIICRPLRARVLTTRRRVRLIALVLAVALLSAGPFLVGV 180
 Db 121 CTYATLLHMTALSVERYLAIICRPLRARVLTTRRRVRLIALVLAVALLSAGPFLVGV 180
 Qy 181 QDQGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAAALFSRECRSPA 240
 Db 181 QDQGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAAALFSRECRSPA 240
 Qy 241 QLGALRVMLWVTTAYFFLPFLCLSLYGLIGRELWSSRRPLRGPAASGRGRHQTQVRL 300
 Db 241 QLGALRVMLWVTTAYFFLPFLCLSLYGLIGRELWSSRRPLRGPAASGRGRHQTQVRL 300
 Qy 301 LVVLAFLIICWLPFHVGRIIYINTEDSRMWFYQYFNIVALQFLYLSASINPILYNLISK 360
 Db 301 LVVLAFLIICWLPFHVGRIIYINTEDSRMWFYQYFNIVALQFLYLSASINPILYNLISK 360
 Qy 361 KYRAAFAKLLARKSRPRGHRSDTAGVAGDTGDTGVGTYSANVKTWG 412
 Db 361 KYRAAFAKLLARKSRPRGHRSDTAGVAGDTGDTGVGTYSANVKTWG 412

RESULT 8
 ID ABB09535
 AC ABB09535 standard; protein; 412 AA.

XX
 AC ABB09535;
 XX
 DT 22-OCT-2002 (first entry)
 XX
 DE Human motilin receptor.
 XX

KW Human; motilin receptor; appetite; food intake; agonist; analogue;
 KW undernutrition; anorexia; cachexia; malignant disease; infection;
 KW inflammatory disease; weight loss; antagonist; obesity; anorectic;
 KW anabolic; ghrelin receptor homologue; receptor.
 XX
 OS Homo sapiens.
 XX
 PN WO200260472-A1.
 XX
 PD 08-AUG-2002.
 XX
 PF 31-JAN-2002; 2002WO-JP00765.
 XX
 PR 31-JAN-2001; 2001JP-0024423.
 XX
 PA (CHUS) CHUGAI SEIYAKU KK.
 XX
 PI Inui A, Asakawa A, Kaga T;
 XX
 DR WPI; 2002-619206/66.
 XX
 PT Remedies for diseases with hypo-nutrition status e.g. inappetence and
 PT cachexia, containing ghrelin or its analogs including agonists and
 PT antagonists -
 XX
 PS Disclosure; Fig 1B; 50pp; Japanese.
 XX
 CC The invention relates to the use of ghrelin or its analogues for the
 CC treatment of diseases associated with undernutrition such as anorexia,
 CC and also relates to the use of ghrelin antagonists for the prevention
 CC or treatment of obesity. The invention additionally discloses a method
 CC for screening ghrelin agonists or antagonists by measuring the amount
 CC of food intake, neuropeptide Y (NPY) expression, binding of NPY to
 CC NPY receptor Y1, oxygen consumption, gastric emptying, or activity of
 CC the vagus nerve. Intracerebroventricular (ICV) administration of ghrelin
 CC in animals was found to increase food intake over a period of 24 hours.
 CC Ghrelin and its analogues may therefore be used to treat conditions
 CC such as loss of appetite, anorexia, cachexia, malignant diseases, and
 CC weight loss associated with infection or inflammatory diseases.
 CC Conversely, ghrelin antagonists may be used in the treatment of obesity.
 CC The present sequence represents the human motilin receptor, a homologue
 CC of the ghrelin receptor (ABB09534) which is referred to in the disclosure
 CC of the invention.
 XX
 SQ Sequence 412 AA;

Query Match 99.4%; Score 2141; DB 23; Length 412;
 Best Local Similarity 99.5%; Pred. No. 2.5e-193;
 Matches 410; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 MGSPWNGSDGPEGAREPPWPPALPPCDERRCSPPPLGALVPVTAACLVFVVGSGNVTV 60
 Db 1 MGSPWNGSDGPEGAREPPWPPALPPCDERRCSPPPLGALVPVTAACLVFVVGSGNVTV 60
 Qy 61 MLIGRYDRMTTNNLYLGSMAVSDLLILLGLPPDLYRLWRSRPPWFGPLCRLSLYVGE 120
 Db 61 MLIGRYDRMTTNNLYLGSMAVSDLLILLGLPPDLYRLWRSRPPWFGPLCRLSLYVGE 120
 Qy 121 CTYATLLHMTALSVERYLAIICRPLRARVLTTRRRVRLIALVLAVALLSAGPFLVGV 180
 Db 121 CTYATLLHMTALSVERYLAIICRPLRARVLTTRRRVRLIALVLAVALLSAGPFLVGV 180
 Qy 181 QDQGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAAALFSRECRSPA 240
 Db 181 QDQGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAAALFSRECRSPA 240
 Qy 241 QLGALRVMLWVTTAYFFLPFLCLSLYGLIGRELWSSRRPLRGPAASGRGRHQTQVRL 300
 Db 241 QLGALRVMLWVTTAYFFLPFLCLSLYGLIGRELWSSRRPLRGPAASGRGRHQTQVRL 300
 Qy 301 LVVLAFLIICWLPFHVGRIIYINTEDSRMWFYQYFNIVALQFLYLSASINPILYNLISK 360
 Db 301 LVVLAFLIICWLPFHVGRIIYINTEDSRMWFYQYFNIVALQFLYLSASINPILYNLISK 360

QY 361 KYRAAFKLLARKSRPRGFRHSDRTAGEVAGDTGGDTGVTYTSANVKTWG 412
 DB 361 KYRAAFKLLARKSRPRGFRHSDRTAGEVAGDTGGDTGVTYTSANVKTWG 412

RESULT 9
 AAB68477
 ID AAB68477 standard; Protein; 400 AA.
 AC AAB68477;
 XX
 XX
 DT 23-JUL-2001 (first entry)
 XX
 XX Amino acid sequence of a rabbit motilin receptor polypeptide.
 DE
 XX Motilin receptor; gastrointestinal disease; gastric motility disorder;
 KW gastroparesis; irritable bowel syndrome; diarrhoea.
 XX
 XX Oryctolagus cuniculus.
 OS
 XX
 PN WO200132710-A1.
 XX
 XX 10-MAY-2001.
 PD
 XX 25-OCT-2000; 2000WO-US29426.
 PF
 XX 29-OCT-1999; 99US-0162264.
 PR
 XX (MERI) MERCK & CO INC.
 PA
 XX Tan C, McKee K;
 PI
 XX WPI; 2001-343479/36.
 DR
 XX N-PSDB; AAF95448.
 DR
 XX Novel polypeptides related to dog and rabbit motilin receptor
 PT polypeptide, comprising unique regions from dog and motilin receptor
 PT amino acid sequence, useful for identifying compounds for treating
 PT diarrhoea in humans
 XX
 XX Claim 8; Page 18; 42pp; English.
 PS
 XX The present sequence represents a rabbit motilin receptor polypeptide.
 CC The specification describes a unique sequence present in exon 1 of
 CC the dog motilin receptor, which is not present in human or Sphaeroides
 CC nepheus 75E7 motilin receptor sequences. The unique nucleic acid
 CC sequence is useful for measuring the ability of a compound to affect
 CC motilin receptor activity. Motilin receptor polynucleotides and
 CC polypeptides are used to identify therapeutic compounds which are
 CC useful for treating gastrointestinal diseases and disorders such as
 CC gastric motility disorders, gastroparesis, irritable bowel syndrome,
 CC and diarrhoea.
 XX
 SQ Sequence 400 AA;
 Query Match 81.1%; Score 1747; DB 22; Length 400;
 Best Local Similarity 83.9%; Pred. No. 3.5e-156;
 Matches 344; Conservative 11; Mismatches 43; Indels 12; Gaps 3;

QY 1 MGSPWNGDGPGEAREPPWALPPCDERRCSPFFPLGALVPVTAIVCLCLFVGVSGNVVTV 60
 DB 1 MGSPWNGDGPGEAREPPWALPPCDERRCSPFFPLGTLVPVTAIVCLGLFVGVSGNVVTV 60
 QY 61 MLIGRYDMRTTNYLIGSMAVSDLLILGLPFDLYRLWRSRPVFPGLCRLSLYVGE 120
 DB 61 LLIGRYDMRTTNYLIGSMAVSDLLILGLPFDLYRLWRSRPVFPGLCRLSLYVGE 120
 QY 121 CTYATLLHMTALSVRYLAICRPARVLTTRRVRVLAIVLWAVALLSAGPFLVGV 180
 DB 121 CTYASLHMTALSVRYLAICRPARVLTTRRVRVLAIVLWAVALLSAGPFLVGV 180
 QY 181 QDPGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPETAEEAALFSPRECRPSA 240

DB 181 QDPAVFAAPDRNGTV-----PLDPSP-----APASPPSGP-GAEEAALFSPRECRPSA 228
 QY 241 QLGALRVMLWVTYAYFPFLPCLISILYGLTGRELWSRRPLRGPAASGRGRHQTVRVL 300
 DB 229 QLGALLRVMLWVTYAYFPFLPCLISILYGLTARQWGRGRLGCPAATGRGRHQTVRVL 288
 QY 301 LVVLAFLIICWLPFHVGRIIYINTDSRMVFSQYFNIVALQFLYLSASINPILYNLISK 360
 DB 289 LVVLAFLIICWLPFHVGRIIYINTDSRMVFSQYFNIVALQFLYLSASINPILYNLISK 348
 QY 361 KYRAAFKLLARKSRPRGFRHSDRTAGEVAGDTGGDTGVTYTSANVKT 410
 DB 349 KYRAAARLLRBSRAGPSGCGSRPEQDVAGDTGGDTAGCTTTSANTKT 398

RESULT 10
 AAY54146
 ID AAY54146 standard; Protein; 386 AA.
 XX
 AC AAY54146;
 XX
 DT 27-MAR-2000 (first entry)
 XX
 DE Amino acid sequence of the motilin receptor splice variant MTL-R1B.
 XX
 KW Motilin receptor; G-protein coupled receptor; MTL-R1; GPR38;
 KW spliced form; MTL-R1A; MTL-R1B; gastric motility disorder;
 KW functional defect; neurological disorder; scleroderma; colonoscopy;
 KW paraneoplastic syndrome; radiation induced dysmotility; diabetes;
 KW infection; stress-related motility disorder; psychogenic disorder;
 KW gastroparesis; gastro-oesophageal reflux disease; constipation;
 KW chronic idiopathic pseudo obstruction; acute faecal impaction;
 KW postoperative ileus; gallstones; infantile colic; diarrhoea;
 KW irritable bowel syndrome; non-ulcer dyspepsia; non-cardiac chest pain;
 KW endoscopy; duodenal intubation.
 XX
 OS Homo sapiens.
 XX
 PN WO9964436-A1.
 PD 16-DEC-1999.
 XX
 XX 08-JUN-1999; 99WO-US12773.
 PR 12-JUN-1998; 98US-0089098.
 XX (MERI) MERCK & CO INC.
 XX Feighner SD, Patchett AA, Tan C, McKee K, MacNeil D, Howard AD;
 PI Pong S, Smith RG;
 XX WPI; 2000-105868/09.
 DR N-PSDB; AAZ45404.
 XX
 PT Novel receptor protein for screening compounds used in treating
 PT irritable bowel syndrome, constipation and other gastric conditions
 XX
 PS Claim 5; Fig 5; 44pp; English.
 XX

The present sequence represents splice variant MTL-R1B of the motilin receptor. The gene encodes a G-protein coupled receptor, and is designated MTL-R1 (also GPR38). Two spliced forms of GPR38 exist, MTL-R1A (see AAY54145) and MTL-R1B (see AAY54146). MTL-R1A is a functional seven transmembrane domain form, and MTL-R1B is a truncated agonists and antagonists which can be used for treating gastric motility disorders, functional defects, disorders secondary to neurological disorders e.g. scleroderma, paraneoplastic syndromes radiation induced dysmotility, diabetes, infections, stress-related motility disorders, psychogenic disorders, gastroparesis, gastro-oesophageal reflux disease, constipation, chronic idiopathic pseudo obstruction, acute faecal impaction, postoperative ileus, gallstones, infantile colic, irritable

CC bowel syndrome, non-ulcer dyspepsion, non-cardiac chest pain and
 CC diarrhoea. They can also be used in the preparation for colonoscopy,
 CC endoscopy and duodenal intubation.
 XX
 SQ Sequence 386 AA;
 Query Match 73.4%; Score 1581; DB 21; Length 386;
 Best Local Similarity 100.0%; Pred. No. 1.5e-140;
 Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MGSPWNGSDGEGAREPPWALPCDERRCSPFPLGALVPTAVCLCLFVVGVSNGVTV 60
 Db 1 MGSPWNGSDGEGAREPPWALPCDERRCSPFPLGALVPTAVCLCLFVVGVSNGVTV 60
 Qy 61 MLIGRYDRMTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWFGLPCLRLSLYVGE 120
 Db 61 MLIGRYDRMTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWFGLPCLRLSLYVGE 120
 Qy 121 CTYATLLHMTALSVERYLAICRPLRARVLVTRRRVRALIAVLWAVALLSAGPFLVLGVE 180
 Db 121 CTYATLLHMTALSVERYLAICRPLRARVLVTRRRVRALIAVLWAVALLSAGPFLVLGVE 180
 Qy 181 QDPGISVVPGLNGTARIASSPLASPPPLWLSRAPPPSPGPTAEAAALFSRECRPSA 240
 Db 181 QDPGISVVPGLNGTARIASSPLASPPPLWLSRAPPPSPGPTAEAAALFSRECRPSA 240
 Qy 241 QLGLARVLMVWTTAYFFLPFLCLSLYGLIGRELWSSRRPLRGPAAASGRGRHQTVRVL 300
 Db 241 QLGLARVLMVWTTAYFFLPFLCLSLYGLIGRELWSSRRPLRGPAAASGRGRHQTVRVL 300
 RESULT 11
 AAB62653
 ID AAB62653 standard; Protein; 386 AA.
 XX
 AC AAB62653;
 XX
 DT 23-JUL-2001 (first entry)
 XX
 DE Short form of motilin receptor, GPR-38B isoform.
 XX
 KW zsig33; signal transduction; hormone; enzyme; neural development;
 KW gastric contractility; nutrient uptake; digestive; pancreatic; human;
 KW insulin-like growth factor-I; growth hormone; bone; gastrointestinal;
 KW glucose; osteopathic; anorectic; vulnery; immunomodulator; GHS-R;
 KW G-protein coupled receptor; motilin receptor; GPR38; GPR-38B; isoform.
 XX
 OS Homo sapiens.
 XX
 PN WO200138355-A2.
 XX
 PD 31-MAY-2001.
 XX
 PF 22-NOV-2000; 2000WO-US32074.
 XX
 PR 22-NOV-1999; 99US-0166765.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Sheppard PO, Jaspers SR, Deisher TA, Bishop PD;
 XX
 DR WPI; 2001-355879/37.
 XX
 DR N-PSDB; AAF83684.
 XX
 PT Forming reversible peptide receptor complex for purifying cell and
 PT peptides, stimulating signal transduction and modulating hormone
 PT secretion, involves contacting a receptor with zsig33 polypeptide -
 XX
 PS Disclosure; Page 106-109; 11pp; English.
 CC The invention relates to a method of forming a reversible peptide-
 CC receptor complex that involves providing an immobilized receptor, and
 CC contacting the receptor with a zsig33 peptide (comprising residues 24-37

CC of AAB62649), where the receptor binds to the zsig33 peptide. The method
 CC is useful for purifying cells, purifying a peptide, stimulating signal
 CC transduction in a cell expressing a receptor. It is also useful for
 CC modulating secretion of hormones, neural development and/or utilization,
 CC gastric contractility, nutrient uptake, secretion of digestive and
 CC pancreatic enzymes and hormones, secretion of insulin-like growth factor
 CC -I, secretion of non-zsig33 proteins. It is useful for modulating growth
 CC hormone secretion in a mammal having a disease associated with abnormal
 CC levels of growth hormone, such as osteoporosis, bone repair, bone
 CC remodeling, low osteoblast levels, cartilage repair and remodeling,
 CC skeletal dysplasia, immune suppression, obesity, growth retardation,
 CC protein catabolic responses after surgery, cachexia, protein loss,
 CC dwarfism, wound healing and ovulation induction, treating a mammal having
 CC a metabolic disorder requiring neurological feedback, such as satiety
 CC regulation, glucose absorption and metabolism and neuropathy-associated
 CC gastrointestinal disorders, and stimulating glucose-induced insulin
 CC release in a mammal. The present sequence represents the short form of
 CC motilin receptor, GPR-38B (one of the two isoforms of GPR38 which result
 CC from alternative splicing). GPR38 has homology to the human G-protein
 CC coupled receptor, GHS-R.
 XX
 SQ Sequence 386 AA;
 Query Match 73.4%; Score 1581; DB 22; Length 386;
 Best Local Similarity 100.0%; Pred. No. 1.5e-140;
 Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MGSPWNGSDGEGAREPPWALPCDERRCSPFPLGALVPTAVCLCLFVVGVSNGVTV 60
 Db 1 MGSPWNGSDGEGAREPPWALPCDERRCSPFPLGALVPTAVCLCLFVVGVSNGVTV 60
 Qy 61 MLIGRYDRMTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWFGLPCLRLSLYVGE 120
 Db 61 MLIGRYDRMTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWFGLPCLRLSLYVGE 120
 Qy 121 CTYATLLHMTALSVERYLAICRPLRARVLVTRRRVRALIAVLWAVALLSAGPFLVLGVE 180
 Db 121 CTYATLLHMTALSVERYLAICRPLRARVLVTRRRVRALIAVLWAVALLSAGPFLVLGVE 180
 Qy 181 QDPGISVVPGLNGTARIASSPLASPPPLWLSRAPPPSPGPTAEAAALFSRECRPSA 240
 Db 181 QDPGISVVPGLNGTARIASSPLASPPPLWLSRAPPPSPGPTAEAAALFSRECRPSA 240
 Qy 241 QLGLARVLMVWTTAYFFLPFLCLSLYGLIGRELWSSRRPLRGPAAASGRGRHQTVRVL 300
 Db 241 QLGLARVLMVWTTAYFFLPFLCLSLYGLIGRELWSSRRPLRGPAAASGRGRHQTVRVL 300
 RESULT 12
 AAB68476
 ID AAB68476 standard; Protein; 271 AA.
 XX
 AC AAB68476;
 XX
 DT 23-JUL-2001 (first entry)
 XX
 DE Amino acid sequence of a dog motilin receptor exon 1.
 XX
 KW Motilin receptor; gastrointestinal disease; gastric motility disorder;
 KW gastroparesis; irritable bowel syndrome; diarrhoea.
 XX
 OS Canis sp.
 XX
 PN WO200132710-A1.
 XX
 PD 10-MAY-2001.
 XX
 PF 25-OCT-2000; 2000WO-US29426.
 XX
 PR 23-OCT-1999; 99US-0162264.
 XX
 PA (MERI) MERCK & CO INC.
 XX

Db	196	MWVSSVFFFLPVFCLTVLGLIGKLRGRGDTAGGASLRQSHRQTVKMLAVVVFAP	255
Qy	308	IICWLPFHVGRIIYINTEDS---RMVFSQYFNIVALQLFYLSASINPILYNLISKYRA	364
Db	256	IFCWLPHVGRYLFSPGSLFIAGISQYCNLVSVFLFYLSAANPILYNIMSKYRV	315
Qy	365	AAFKLLARKSRPRGFRHSRDTAGEVAGDTGGDTVGYTETSAN	407
Db	316	AVFKLGFEPFSQRKSLTKDESSR-----AWTESSIN	348

Search completed: January 1, 2004, 06:46:43
Job time : 54.1454 Secs

THIS PAGE BLANK (uspro)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 1, 2004, 06:22:35 ; Search time 17.5539 Seconds
(without alignments)
993.061 Million cell updates/sec

Title: US-09-719-485-3
Perfect score: 2155
Sequence: 1 MGSPWNGSDGPEGAREPPWP.....DTGCDTVGTETSANVKTWG 412

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A COMB.pap.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pap.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pap.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pap.*
5: /cgn2_6/ptodata/1/iaa/PTUS COMB.pap.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	863.5	40.1	366	3	US-09-077-675A-13
2	863.5	40.1	366	4	US-09-077-674-13
3	863.5	40.1	366	4	US-09-170-496D-88
4	863	40.0	353	3	US-09-077-675A-3
5	863	40.0	353	4	US-09-077-674-3
6	861.5	40.0	364	3	US-09-077-675A-16
7	861.5	40.0	364	4	US-09-077-674-16
8	860	39.9	361	3	US-09-077-675A-8
9	860	39.9	361	4	US-09-077-674-8
10	857.5	39.8	366	4	US-09-170-496D-210
11	775.5	36.0	302	3	US-09-077-675A-2
12	775.5	36.0	302	4	US-09-077-674-2
13	774.5	35.9	302	3	US-09-077-675A-7
14	774.5	35.9	302	4	US-09-077-674-7
15	674.5	31.3	271	3	US-09-077-675A-12
16	674.5	31.3	271	4	US-09-077-674-12
17	644.5	29.9	289	3	US-09-077-675A-10
18	644.5	29.9	289	4	US-09-077-674-10
19	633	29.4	289	3	US-09-077-675A-5
20	633	29.4	289	4	US-09-077-674-5
21	485.5	22.5	403	4	US-09-170-496D-114
22	483.5	22.4	403	4	US-09-170-496D-224
23	479	22.2	415	4	US-09-545-944-2
24	466	21.6	353	1	US-08-118-270-45
25	466	21.6	353	5	PCT-US93-08528-45
26	413	19.2	410	3	US-08-858-876A-2
27	413	19.2	410	3	US-09-472-880-2

Query Match 40.1%; Score 863.5; DB 3; Length 366;
Best Local Similarity 44.7%; Pred. No. 1.6e-66;

Sequence 4, Appli
Sequence 4, Appli
Sequence 2, Appli
Sequence 4, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 10, Appli
Sequence 10, Appli
Sequence 8, Appli
Sequence 6, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-077-675A-13
; Sequence 13, Application US/09077675A
; Patent No. 6242199
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 366 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-077-675A-13

Sequence 13, Application US/09077675A

Patent No. 6242199

GENERAL INFORMATION:

APPLICANT: Pai, Lee-Yuh

APPLICANT: Feighner, Scott C.

APPLICANT: Howard, Andrew D.

APPLICANT: Pong, Sheng-Shung

APPLICANT: Van Der Ploeg, Leonardus H.T.

TITLE OF INVENTION: RECEPTOR ASSAY

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Merck & Co., Inc.

STREET: P.O. Box 2000, 126 E. Lincoln Ave.

CITY: Rahway

STATE: NJ

COUNTRY: USA

ZIP: 07065-0900

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/077,675A

FILING DATE: 3-JUN-1998

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Cocuzzo, Anna L.

REGISTRATION NUMBER: 42,452

REFERENCE/DOCKET NUMBER: 19590P

TELECOMMUNICATION INFORMATION:

TELEPHONE: 732-594-1273

TELEFAX: 732-594-4720

TELEX:

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 366 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-077-675A-13

Matches 186; Conservative 59; Mismatches 106; Indels 65; Gaps 9;

QY 5 WNGSDGPEGA-----REPPWALPPCD---ERRCSPPPLGALVPVTAVCICLCLFVWVSGN 56
 DB 2 WNATPSEEPGNLTADLDWDASPGNDSLGDELLQLFPAPLLAGVATATCVALFVVGIA 61
 QY 57 VVTVMILIGRYDMRTTNNLYLGSMAVSDLLILGLPDPDLVRLWRSRPWVFGPLLCRLSL 116
 DB 62 LITMLVVSFRFRELRTTNNLYLSSMAFSDLLIFLCMPDLVRLWQYRPMWFGDILLCKLFQ 121
 QY 117 VGEQCTVATLHMTALSVRYLAICRPLARVLVTRRRVRLAVLWAVALLSAGPPLFL 176
 DB 122 VSECTVATLITLALSVRYFAICFPLRAKVVVTKGRVKLVFIWAVAFCSAGPTFVL 181
 QY 177 VGVQDPGISVVPGLNGTARIASSPLASSPLMLSRAPPSPPSGPTAEAAALFSGREC 236
 DB 182 VGVHE-----NGT-----DP-W-----DINECR 199
 QY 237 PS--PAQALGALRVMLVTTAYFPFLPCLSLILGLIGRELWSSRRPLRGPASGRGRGHR 294
 DB 200 PTEFAVRSGLLTVMWVSSIFFPFLPVCLVLYSLIGRLWRRRGDAVVGASLRDQNHK 259
 QY 295 QTVRVLLVVLAFIICWLPFHVGRITIIYNTEDS---RMMVFQYFNIVALQLFVLSASIN 351
 DB 260 QTVKMLAVVFAFICWLPFHVGRYLFKSFPESGLEIAQISQYCNLVSVFLVLSA 319
 QY 352 PILYNIISKYRAAFKLLARKSRPGRHRSRDTAGEVAGDTGGDTGVGTETSAN 407
 DB 320 PILYNIIMSKYRVAVFRLGFPESQKSLTKDESSR-----AWTESSIN 365

RESULT 2

US-09-077-674-13
 ; Sequence 13, Application US/09077674
 ; Patent No. 6531314
 ; GENERAL INFORMATION:
 ; APPLICANT: Arena, Joseph P.
 ; APPLICANT: Cully, Doris F.
 ; APPLICANT: Feighner, Scott D.
 ; APPLICANT: Howard, Andrew D.
 ; APPLICANT: Liberator, Paul A.
 ; APPLICANT: Schaeffer, James M.
 ; APPLICANT: Van Der Ploeg, Leonardus
 ; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR FAMILY
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESSES:
 ; ADDRESSEE: Merck & Co., Inc.
 ; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
 ; CITY: Rahway
 ; STATE: NJ
 ; COUNTRY: USA
 ; ZIP: 07065-0900
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/077, 674
 ; FILING DATE: 3-JUN-1998
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Cocuzzo, Anna L.
 ; REGISTRATION NUMBER: 42,452
 ; REFERENCE/DOCKET NUMBER: 19589P
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 732-594-1273
 ; TELEFAX: 732-594-4720
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:
 LENGTH: 366 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-077-674-13

Query Match 40.1%; Score 863.5; DB 4; Length 366;
 Best Local Similarity 44.7%; Pred. No. 1.6e-66;
 Matches 186; Conservative 59; Mismatches 106; Indels 65; Gaps 9;

QY 5 WNGSDGPEGA-----REPPWALPPCD---ERRCSPPPLGALVPVTAVCICLCLFVWVSGN 56
 DB 2 WNATPSEEPGNLTADLDWDASPGNDSLGDELLQLFPAPLLAGVATATCVALFVVGIA 61
 QY 57 VVTVMILIGRYDMRTTNNLYLGSMAVSDLLILGLPDPDLVRLWRSRPWVFGPLLCRLSL 116
 DB 62 LITMLVVSFRFRELRTTNNLYLSSMAFSDLLIFLCMPDLVRLWQYRPMWFGDILLCKLFQ 121
 QY 117 VGEQCTVATLHMTALSVRYLAICRPLARVLVTRRRVRLAVLWAVALLSAGPPLFL 176
 DB 122 VSECTVATLITLALSVRYFAICFPLRAKVVVTKGRVKLVFIWAVAFCSAGPTFVL 181
 QY 177 VGVQDPGISVVPGLNGTARIASSPLASSPLMLSRAPPSPPSGPTAEAAALFSGREC 236
 DB 182 VGVHE-----NGT-----DP-W-----DINECR 199
 QY 237 PS--PAQALGALRVMLVTTAYFPFLPCLSLILGLIGRELWSSRRPLRGPASGRGRGHR 294
 DB 200 PTEFAVRSGLLTVMWVSSIFFPFLPVCLVLYSLIGRLWRRRGDAVVGASLRDQNHK 259
 QY 295 QTVRVLLVVLAFIICWLPFHVGRITIIYNTEDS---RMMVFQYFNIVALQLFVLSASIN 351
 DB 260 QTVKMLAVVFAFICWLPFHVGRYLFKSFPESGLEIAQISQYCNLVSVFLVLSA 319
 QY 352 PILYNIISKYRAAFKLLARKSRPGRHRSRDTAGEVAGDTGGDTGVGTETSAN 407
 DB 320 PILYNIIMSKYRVAVFRLGFPESQKSLTKDESSR-----AWTESSIN 365

RESULT 3

US-09-170-496D-88
 ; Sequence 88, Application US/09170496D
 ; Patent No. 6553339
 ; GENERAL INFORMATION:
 ; APPLICANT: Behan, Dominic P.
 ; APPLICANT: Chalmers, Derek T.
 ; APPLICANT: Liaw, Chen W.
 ; TITLE OF INVENTION: Receptors
 ; FILE REFERENCE: AREN-0040
 ; CURRENT APPLICATION NUMBER: US/09/170,496D
 ; CURRENT FILING DATE: 1998-10-13
 ; NUMBER OF SEQ ID NOS: 294
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 88
 ; LENGTH: 366
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-170-496D-88

Query Match 40.1%; Score 863.5; DB 4; Length 366;
 Best Local Similarity 44.7%; Pred. No. 1.6e-66;
 Matches 186; Conservative 59; Mismatches 106; Indels 65; Gaps 9;

QY 5 WNGSDGPEGA-----REPPWALPPCD---ERRCSPPPLGALVPVTAVCICLCLFVWVSGN 56
 DB 2 WNATPSEEPGNLTADLDWDASPGNDSLGDELLQLFPAPLLAGVATATCVALFVVGIA 61
 QY 57 VVTVMILIGRYDMRTTNNLYLGSMAVSDLLILGLPDPDLVRLWRSRPWVFGPLLCRLSL 116
 DB 62 LITMLVVSFRFRELRTTNNLYLSSMAFSDLLIFLCMPDLVRLWQYRPMWFGDILLCKLFQ 121

QY 117 VGGCTVATLHMTALSVRYLAICRPLRVLVTRRRVRLIAVLWAVALLSAGPFLFL 176
DB 122 VSSCTVATLHMTALSVRYFAICFPLRAKVVVTKGRVLFVIMAVAFCSAGPIFVL 181
QY 177 VGGVODPGISVVPGLNCTARIASSPLASSPLMLSRAPPSPGPGTAAALFSRECR 236
DB 182 VGVHE-----NGT-----DP-W-----DNECR 199
QY 237 PS--PAQLGALRVMLVTTTAYFFLPCLSLYLGLIGRELWSSRRPLRGPAASGRERGRH 294
DB 200 PTEFAVRSGLLTVMWVSSIFFLPVCLVLSLIGRLWRRRGDAVVGASLRDNHK 259
QY 295 QTVRVLLVVLVLAFLICWLPFHVGRIIYNTEDS---RMMYFSQYFNIVALQLFYLSASIN 351
DB 260 QTVKMLAVVVFATILCWLPFHVGRIYFSKSPGSLGIAQISQYCNLVSFVLYLSAAIN 319
QY 352 PILYNLSKKYRAAFAKLLARKSRPGRPHRSRDTAGEVAGDTGDTGVGTETSAN 407
DB 320 PILYNMSKKYRVAVFRLGFEPPSQRLSTLKDSSR-----AWTESSIN 365

RESULT 4

US-09-77-675A-3

; Sequence 3, Application US/09077675A
; Patent No. 6242199
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 353 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-09-077-675A-3
Query Match
Best Local Similarity 40.0%; Score 863; DB 3; Length 353;
Matches 186; Conservative 53; Mismatches 95; Indels 66; Gaps 9;

QY 19 WPALPPCD---ERRCSPPFPLGALVPTAVCLCLFVVVGSGNVVTVMLIGRYDRMTITNL 75
DB 8 WDAPPENDSLVEELLPLFPPTPLAGVTATCVALFVVGIAGNLLTMLVWSRPREMRTITNL 67
QY 76 YLGSMAVSDLLILLGLPFDLYRLWRSPVWPGPLLCRLSLYVCEGCTYATLLHMTALSVE 135
DB 68 YLSSMAFSDLLIFLCMPLDLFLWQYRPMNLGNLCKLFQFVBSSECTYATVLTITALSVE 127
QY 136 RYLAICRPLRVLVTRRRVRLIAVLWAVALLSAGPFLVGVGEODPGISVVPGLNGTA 195
DB 128 RYFAICFPLRAKVVVTKGRVLFVIMAVAFCSAGPIFVLVGEHD-----NGT- 177
QY 196 RIASSPLASSPLMLSRAPPSPGPGTAAALFSRECRPS--PAQLGALRVMLWVTT 253
DB 178 -----DPRD-----TNECRATEFAVRSGLLTVMWVSS 205
QY 254 AVFPLPCLSLYLGLIGRELWSSRRPLRGPAASG---RERGHROTVRVLLVVLVLAFLIC 310
DB 206 VFFPLFVCLTVLSYSLGRKLM---RRKRGAAVGSLSRDNHKTQVLMVAVVVFALFIC 262
QY 311 WLPFHVGRIIY---INTDSRMMYFSQYFNIVALQLFYLSASINPILYNLSKKYRAAFA 367
DB 263 WLPFHVGRIYFSKSLSPGSLGIAQISQYCNLVSFVLYLSAAINPILYNMSKKYRVAVP 322
QY 368 KLLARKSRPGRPHRSRDTAGEVAGDTGDTGVGTETSAN 407
DB 323 KLIGPPFPQRLSTLKDSSR-----AWTESSIN 352

RESULT 5

US-09-077-674-3

; Sequence 3, Application US/09077674
; Patent No. 6531314
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR FAMILY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,674
; FILING DATE: 3-JUN-1998
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19589P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:

```
; LENGTH: 353 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-077-674-3

Query Match 40.0%; Score 863; DB 4; Length 353;
Best Local Similarity 46.5%; Pred. No. 1.7e-66;
Matches 186; Conservative 53; Mismatches 95; Indels 66; Gaps 9;

QY 19 WPALPPCD---ERRCSPFPLGALVPVTAVALCLFVGVSGNVVTLIGRYDMRTTNL 75
DB 8 WDAPPENDSLVEELLPLFPPTLLAGVTATCVAFVVGAGNLLTLMVSVRFEMRTTNL 67

QY 76 YLGSMVSDLLIILGLPFDLYLRWRSPWPVFGPILCLSLYVGEGETYATLLHMTLSVE 135
DB 68 YLSSMAFSDLLIFLCMPDLFLRLQVYRPMNLGNLLCKLFQVSESCYATVLTITLSVE 127

QY 136 RYLAICRPLARVLTTRRRVRLIAVLWAVALLSAGPFLFLVGEODPGISVVVPGNLGTA 195
DB 128 RYFAICFPLRAKVVTGRVKLVILWVAVAFCSAGPIFVLVGEHD-----NGT- 177

QY 196 RIASSPLASSPPLWLSRAPPPSPGPGTAEAAALFSRECRPS--PAQLGALRVMLVTT 253
DB 178 -----DPRD-----TNECRATEFAVRSGLLTVMVWSS 205

QY 254 AYFPLFPLCLSIYGLIGRLWSSRPLRGPAAAG---RRGRQTVRVLLVVLAFIIC 310
DB 206 VFFPLFVFCVLTVLSLIGRLW---RRKRGAAVGSRLRDQNHQTVKMLAVVVFILC 262

QY 311 WLPFHVGRIIY---INTEDSRMYFQYFNIVALQLFYLSASINPILYNLISKYRAAAF 367
DB 263 KLPFHVGRIYLSKSPGSEVIAQISQYCNLVSVFLYSAANPILYINIMSKYRVAVF 322

QY 368 KLLARKSRPRGRPHRRDAGEVAGDTGDTGVYGTETSN 407
DB 323 KLAGEPFPQSKLSTLKDESSR-----AWTESSIN 352

RESULT 6
US-09-077-675A-16
; Sequence 16, Application US/09077675A
; Patent No. 6242199
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
```

```
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 364 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-077-675A-16

Query Match 40.0%; Score 861.5; DB 3; Length 364;
Best Local Similarity 48.3%; Pred. No. 2.4e-66;
Matches 185; Conservative 52; Mismatches 81; Indels 65; Gaps 11;

QY 5 WNGSDGPEGAREP-----PWPALPPCD---ERRCSPFPLGALVPVTAVALCLFVVGVS 55
DB 2 WNAT--PSBEPEPNVTLDLDWDASPGNDSLPLDELLPLFPAPLLAGVTATCVAFVVGIS 59

QY 56 NVVTVMILIGRYDMRTTNLYLGSMAVSDLLIILGLPFDLYLRWRSPWPVFGPILCLSL 115
DB 60 NLLTLMVSVRFELRTTNLYLSSMAFSDLLIFLCMPDLVRLWQYRPMNFGDLLCKLFQ 119

QY 116 YVGEGETYATLLHMTLSVERYLAICRPLARVLTTRRRVRLIAVLWAVALLSAGPFLF 175
DB 120 FVSESCYATVLTITLSVERYFAICFPLRAKVVTGRVKLVILWVAVAFCSAGPIFV 179

QY 176 LVGEODPGISVVVPGNLGTAIASSPLASSPPLWLSRAPPPSPGPGTAEAAALFSREC 235
DB 180 LVGEVHE-----NGT-----DPRD-----TNEC 197

QY 236 RPS--PAQLGALRVMLVTTAVFPLFPLCLSIYGLIGRLWSSRPLRGPAAAG---RE 290
DB 198 RATEFAVRSGLLTVMVWSSVFFFLVFCVLTVLSLIGRLW--RR--RCDAAVAGSLRD 253

QY 291 RGRQTVRVLLVVLAFIICWLPFHVGRIIYINTEDS---RMYFQYFNIVALQLFYLS 347
DB 254 QNHQTVKMLAVVVFILCWLPHFVGRVLFSSPFGSLIEIAQISQYCNLVSVFLYLS 313

QY 348 ASINPILYNLISKYRAAFKLL 370
DB 314 AAINPILYINIMSKYRVAVFVFKLL 336

RESULT 7
US-09-077-674-16
; Sequence 16, Application US/09077674
; Patent No. 6531314
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR FAMILY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
```



```

SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/077,674
FILING DATE: 3-JUN-1998
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19589P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 364 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-077-674-16

Query Match 40.0%; Score 861.5; DB 4; Length 364;
Best Local Similarity 48.3%; Pred. No. 2.4e-66;
Matches 185; Conservative 52; Mismatches 81; Indels 65; Gaps 11;

Qy 5 WNSDGPGEAREP-----PWPALPPCD---ERRCSPPFGLAVPVTVAVCLCLFVVGVS 55
Db 2 WNAT--PSEBPEPNVTLDDWDASPGNDSLPLDELLPLFPAPLLAGVATATCVAFVVGIS 59

Qy 56 NVVTMLIGRYDMRTTTLNLYLGSMAVSDLLILGLPDLVRLWRSPWVFGPGLCLRLSL 115
Db 60 NLLTMLVVSFRRLTTLNLYLSSMAFSDLLIFLCMPDLVRLWQYRPMWFGDLLCKLFQ 119

Qy 116 YVGECTYATLLHMTALSVERYLAICRPLARVLTTRRRVRLAVLWAVALLSAGPFLF 175
Db 120 FVSECTYATVLTALSVERYFAICPLRAKVVTTRKRVKLVILVWAVAFCSAGPIFV 179

Qy 176 LVGVEQDPGIVSVVPLNGTARIASSPLASSPPLWLRAPPPSPGSGTETAALFRRRC 235
Db 180 LVGVEH-----NGT-----DPRD-----TNEC 197

Qy 236 RPS--PAQLGALRVMLVTTAYFPFLPCLSLYLIGRELWSSRRPLRGPAASG---RE 290
Db 198 RATEFAVRSGLLTVMWVSSVFFLPVCLTVLYSLIGRLW---RR--RGDAVAGSLRD 253

Qy 291 RGHRTQVRVLLVVLAFIICWLPVHVGRIIYINTEDS---RMVFSQYFNIVALQLFYL 347
Db 254 QNHKQTVKMLAVVFAFVLCWLPVHVGRLFSKSPFGSLEIAQISQYCNLVSVLFYLS 313

Qy 348 ASINPILYNIISKYRAAFKLL 370
Db 314 AAINPILYNIIMSKYRVAVFKLL 336

RESULT 8
US-09-077-675A-8
Sequence 8, Application US/09077675A
Patent No. 6242199
GENERAL INFORMATION:
APPLICANT: Pai, Lee-Yuh
APPLICANT: Feighner, Scott C.
APPLICANT: Howard, Andrew D.
APPLICANT: Pong, Sheng-Shung
APPLICANT: Van Der Ploeg, Leonardus H.T.
TITLE OF INVENTION: RECEPTOR ASSAY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.

CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/077,675A
FILING DATE: 3-JUN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19590P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 361 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-077-675A-8

Query Match 39.9%; Score 860; DB 3; Length 361;
Best Local Similarity 46.1%; Pred. No. 3.2e-66;
Matches 183; Conservative 57; Mismatches 97; Indels 60; Gaps 8;

Qy 19 WPALPPCD---ERRCSPPFGLAVPVTVAVCLCLFVVGVSNNVTVMLIGRYDMRTTTL 75
Db 16 WDASPGNDSIGDELLQLFPAPLLAGVATATCVAFVVGIAGNLLTMLVSRPRELRTTTL 75

Qy 76 YLGSMAVSDLLILGLPDLVRLWRSPWVFGPGLCLRLSLYVGECTYATLLHMTALSVE 135
Db 76 YLSSMAFSDLLIFLCMPDLVRLWQYRPMWFGDLLCKLFQVSECTYATVLTALSVE 135

Qy 136 RYLAICRPLARVLTTRRRVRLAVLWAVALLSAGPFLFVGVQDPGIVSVVPLNGTA 195
Db 136 RYFAICPPLRAKVVTTRKRVKLVIFVWAVAFCSAGPIFVLVGVHE-----NGT- 185

Qy 196 RIASSPLASSPPLWLRAPPPSPGSGTETAALFRRRCPS--PAQLGALRVMLVTT 253
Db 186 -----DP--W-----DTNECRTEPAVRSGLLTVMWVSS 213

Qy 254 AYFPLPCLSLYLIGRELWSSRRPLRGPAASGRERGHRTQVRVLLVVLAFIICWLP 313
Db 214 IFPELPVCLTVLYSLIGRLWRRRGDAVVGASLRDQNHKQTVKMLAVVFAFVLCWLP 273

Qy 314 FHVGRITTYINTEDS---RMVFSQYFNIVALQLFYLASINPILYNIISKYRAAFKLL 370
Db 274 FHVGRYLFKSPFGSLEIAQISQYCNLVSVFLYLSAANPILYNIIMSKYRVAVFKLL 333

Qy 371 LARKSRPRGPHRRSDTAGEVAGDTGGDTGYTETSAN 407
Db 334 GFEPFSQRKSLTLKDESSR-----AWTESSIN 360

RESULT 9
US-09-077-674-8
Sequence 8, Application US/09077674
Patent No. 6531314
GENERAL INFORMATION:
APPLICANT: Arena, Joseph P.
APPLICANT: Cully, Doris F.

```

APPLICANT: Feighner, Scott D.
APPLICANT: Howard, Andrew D.
APPLICANT: Liberator, Paul A.
APPLICANT: Schaeffer, James M.
APPLICANT: Van Der Ploeg, Leonardus
TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR FAMILY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/077,674
FILING DATE: 3-JUN-1998
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 195899
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 361 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-077-674-8

Query Match 39.9%; Score 860; DB 4; Length 361;
Best Local Similarity 46.1%; Pred. No. 3-2e-66;
Matches 183; Conservative 57; Mismatches 97; Indels 60; Gaps 8;
QY 19 WPALPPCD---ERRCSFPFPGALVPVAVCLCLFVVGSGNVVTVMLIGRYRDMRTTNL 75
DB 16 WDASPGNDSLGDELLQFPAPLAGVTATCVALFVVGAGNLLTLMVWSRFRRLRTTNL 75
QY 76 YLGSMAVSDLLILGLPFDLYRLWRSPWVPGPLCLRLSYVCEGCTYATLLHMTALSVE 135
DB 76 YLSSMAFSDLLIFLCMPDLVRLWQYRPMWFGDLCCKLFQVSESCYATVLTITALSVE 135
QY 136 RYLAICRPLRVRVTRRRVRLIALVLAVALLSAGFPFLVGVGVDPGISVVPGLNGTA 195
DB 136 RYFAICFPLRAKVVTGRVKLVFIWAVAFSAGPIFVLGVGVEH-----NGT- 185
QY 196 RIASSPLASSPPLWLSRAPPPSPGPETAALFASRCRPS--PAQLGALRVMLVWTT 253
DB 186 -----DP--W-----DTNECRPTEFAVRSGLLTVMVWVSS 213
QY 254 AYFPLFCLISLYLIGRELMSRRLRPGPAASGRGHRQTVRVLLVVLAFICWLP 313
DB 214 IFPLFPVCLTVLSLYLIGRELMSRRLRPGPAASGRGHRQTVRVLLVVLAFICWLP 273
QY 314 FHVGRIIYNTEDS---RMYFSQYFNIVALQLFYLSASINPILYLNLSKKYRAAAFKLL 370
DB 274 FHVGRIYFSKSPFGSLIAQISQYCNLVSVFLYLSAIPILYLNLSKKYRAVAFRL 333
QY 371 LARKSRPRGHRSDTAGVAGDTGGDTVGYTETSAN 407

Db 334 GPEFPQRSKSLTKDESSR-----AWTESSIN 360
RESULT 10
US-09-170-496D-210
Sequence 210, Application US/09170496D
Patent No. 6555339
GENERAL INFORMATION:
APPLICANT: Behan, Dominic P.
APPLICANT: Chalmers, Derek T.
APPLICANT: Liaw, Chen W.
TITLE OF INVENTION: No. 6555339-Endogenous, Constitutively Activated Human G Protein-
FILE REFERENCE: AREN-0040
CURRENT APPLICATION NUMBER: US/09/170,496D
CURRENT FILING DATE: 1998-10-13
NUMBER OF SEQ ID NOS: 294
SOFTWARE: Patent in version 3.1
SEQ ID NO 210
LENGTH: 366
TYPE: PRT
ORGANISM: Homo sapiens
US-09-170-496D-210

Query Match 39.8%; Score 857.5; DB 4; Length 366;
Best Local Similarity 44.5%; Pred. No. 5.3e-66;
Matches 185; Conservative 59; Mismatches 107; Indels 65; Gaps 9;
QY 5 WNGSDGPEGA-----REPPWALPPCD---ERRCSFPFPGALVPVAVCLCLFVVGSGN 56
DB 2 WNATPSEEPFNLTADLDWDASPGNDSLGDELLQFPAPLAGVTATCVALFVVGAGN 61
QY 57 VVTVMILIGRYRDMRTTNLYLGSMAVSDLLILGLPFDLYRLWRSPWVPGPLCLRLSLY 116
DB 62 LUTLWVSRFRRLRTTNLYLSSMAFSDLLIFLCMPDLVRLWQYRPMWFGDLCCKLFQ 121
QY 117 VGEGCTYATLLHMTALSVERYLAICRPLRVRVTRRRVRLIALVLAVALLSAGFPFL 176
DB 122 VSESCYATVLTITALSVERYFAICFPLRAKVVTGRVKLVFIWAVAFSAGPIFVL 181
QY 177 VGVGQDPSVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPETAALFASRCR 236
DB 182 VGVGVEH-----NGT-----DP--W-----DTNECR 199
QY 237 PS--PAQLGALRVMLVWTTAYFPLFCLISLYLIGRELMSRRLRPGPAASGRGHR 294
DB 200 PTEFAVRSGLLTVMVWVSSIFPLFPVCLTVLSLYLIGRELMSRRLRPGPAASGRGHR 259
QY 295 QTVRVLLVVLAFICWLPFHVGRIIYNTEDS---RMYFSQYFNIVALQLFYLSASIN 351
DB 260 QTKKMLAVVVFALICWLPFHVGRIYFSKSPFGSLIAQISQYCNLVSVFLYLSA 319
QY 352 PILYNLSKKYRAAAFKLLLARKSRPRGHRSDTAGVAGDTGGDTVGYTETSAN 407
DB 320 PILYNLSKKYRAVAFRLIGFPGFSPQKSLTKDESSR-----AWTESSIN 365

RESULT 11
US-09-077-675A-2
Sequence 2, Application US/09077675A
Patent No. 6242199
GENERAL INFORMATION:
APPLICANT: Pai, Lee-Yuh
APPLICANT: Feighner, Scott C.
APPLICANT: Howard, Andrew D.
APPLICANT: Pong, Sheng-Shung
APPLICANT: Van Der Ploeg, Leonardus H.T.
TITLE OF INVENTION: RECEPTOR ASSAY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway

```

; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 302 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-077-675A-2

Query Match 36.0%; Score 775.5; DB 3; Length 302;
Best Local Similarity 46.6%; Pred. No. 4.7e-59;
Matches 166; Conservative 48; Mismatches 79; Indels 63; Gaps 8;

QY 60 VMLIGRYDMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSPWVFGPLLCRLSLYVGE 119
Db 1 MLVVSFRFMRRTTNNLYLSSMAFSDLLIFLCMPDLFLRWQYRPMNLNLLCKLQFVSE 60
QY 120 GCTYATLLHMTALSVERYLAICRPLRVLVTRRRVRLTAVLWAVALLSAGPFLVLGV 179
Db 61 SCTYATVLTITALSVERYFAICPLRAKVVVTKGRVKLVILVWVAFCSAGPIFVLGV 120
QY 180 EQDPGISVVPGLNGTARIASSPLASPPPLWLSRAPPPSPGPTAEAAALPSRECRPS - 238
Db 121 EHD-----NGT-----DPRD-----TNECRATE 138
QY 239 -PAQALGALRVLMTVTYAFPLPCLISILYGLIGRELWSSRRPLRGPAAAG---RERGRH 294
Db 139 FAVRSGLLTVMWVSSVFFPLPCLTVLYSLIGRKLW---RRKGEAAVGSLSLRQNHK 195
QY 295 QTVRVLLVVLVLAIIICWLPFHVGRYIY---INTEDSRMMYFSQYFNIVALQLPYLSASIN 351
Db 196 QTVKMLAVVVFAPILCWLPFHVGRYLFKSLSPGVEIAQISQYCNLVSPVLYLSAAIN 255
QY 352 PLYNLISKYRAAFKLLARKSRPRGHRSDTAGVAGDTGGDTGVGTETSAN 407
Db 256 PLYNIMSKKYRVAVFKLLGFEPFSQKLSLTKDESSR-----AWTESSIN 301

RESULT 12
US-09-077-674-2
; Sequence 2, Application US/09077674
; Patent No. 6531314
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus

```

```

; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR FAMILY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,674
; FILING DATE: 3-JUN-1998
; CLASSIFICATION: 536
; PRIOR APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19589P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 302 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-077-674-2

Query Match 36.0%; Score 775.5; DB 4; Length 302;
Best Local Similarity 46.6%; Pred. No. 4.7e-59;
Matches 166; Conservative 48; Mismatches 79; Indels 63; Gaps 8;

QY 60 VMLIGRYDMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSPWVFGPLLCRLSLYVGE 119
Db 1 MLVVSFRFMRRTTNNLYLSSMAFSDLLIFLCMPDLFLRWQYRPMNLNLLCKLQFVSE 60
QY 120 GCTYATLLHMTALSVERYLAICRPLRVLVTRRRVRLTAVLWAVALLSAGPFLVLGV 179
Db 61 SCTYATVLTITALSVERYFAICPLRAKVVVTKGRVKLVILVWVAFCSAGPIFVLGV 120
QY 180 EQDPGISVVPGLNGTARIASSPLASPPPLWLSRAPPPSPGPTAEAAALPSRECRPS - 238
Db 121 EHD-----NGT-----DPRD-----TNECRATE 138
QY 239 -PAQALGALRVLMTVTYAFPLPCLISILYGLIGRELWSSRRPLRGPAAAG---RERGRH 294
Db 139 FAVRSGLLTVMWVSSVFFPLPCLTVLYSLIGRKLW---RRKGEAAVGSLSLRQNHK 195
QY 295 QTVRVLLVVLVLAIIICWLPFHVGRYIY---INTEDSRMMYFSQYFNIVALQLPYLSASIN 351
Db 196 QTVKMLAVVVFAPILCWLPFHVGRYLFKSLSPGVEIAQISQYCNLVSPVLYLSAAIN 255
QY 352 PLYNLISKYRAAFKLLARKSRPRGHRSDTAGVAGDTGGDTGVGTETSAN 407
Db 256 PLYNIMSKKYRVAVFKLLGFEPFSQKLSLTKDESSR-----AWTESSIN 301

RESULT 13
US-09-077-675A-7
; Sequence 7, Application US/09077675A
; Patent No. 6242199
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh

```

APPLICANT: Feighner, Scott C.
APPLICANT: Howard, Andrew D.
APPLICANT: Pong, Sheng-Shung
APPLICANT: Van Der Ploeg, Leonardus H.T.
TITLE OF INVENTION: RECEPTOR ASSAY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS: 16
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA: US/09/077.675A
APPLICATION NUMBER: US/09/077.675A
FILING DATE: 3-JUN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19590P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 302 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-077-675A-7

Query Match	35.9%	Score	774.5;	DB	3;	Length	302;
Best Local Similarity	46.5%	Pred. No.	5.7e-59;				
Matches	164;	Conservative	51;	Mismatches	81;	Indels	57;
Gaps							

QY	60	VMLIGRYDRMTTNNLYGSMVSDLLILGLDPDLVRLWRSRPWFGPPLLCLRLSLVVG	119
DB	1	MLVVSAPREURTTNNLYGSSMAFSDLLIFLCMPDLVLRLWQYRPWNFGDLCKLQFVSE	60
QY	120	GCTYATLLHMTALSVERYLAICRPLRARVLVTRRRVRALIAVLWAVALLSAGPFLFVLGV	179
DB	61	SCIYATVLTITALSVERYFAICRPLRAKVVTGKRVKLVIFVIMAVAFCSAGPIFVLGV	120
QY	180	EQDPGISVVGELNGTARIASSPLASSPPLMLSRAPPPSPGSGPETAEEAALFGRERHPS	238
DB	121	EHE-----NGT-----DP--W-----DTNECRPTE	138
QY	239	-PAQLGALRVMLVWTTAYFELPFLCLILYGLIGRELWSSRRPLRGPAASGRERGHQTV	297
DB	139	FAVRSGLLTVMWVSSIFFELPFLVCLVLSLIGRLWRRRGDAVVGASLRDQNHQTV	198
QY	298	RVLVVVLAFTIICWLPHVGRYIIVNTEDS---RMVFSQYFNIVALQLFVLSASINPIL	354
DB	199	KMLAVVVFARILCWLPHVGRYLFKSFEGSLEIAQISOYCNLVSFVLVLSAANPIL	258
QY	355	YNLISKYRAAAFKLLLRKRRPGRGHRSDRTAGEVAGDTGGDTVGTETTSAN	407
DB	259	YNIMSKYRVAVFLLGLCFEPFSORKLSTLKDSSR-----AMTESIN	301

RESULT 14
US-09-077-674-7

```

; Sequence 7, Application US/09077674
; Patent No. 6531314
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feigner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR FAMILY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,674
; FILING DATE: 3-JUN-1998
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19589P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 302 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-077-674-7

Query Match 35.9%; Score 774.5; DB 4; Length 302;
Best Local Similarity 46.5%; Pred. No. 5,7e-59;
Matches 164; Conservative 51; Mismatches 81; Indels 57;

QY 60 VMLIGRYDMRTNNLYIGSMVSDLLILGLPDLPLRLMRSRPWVFGPLLCRLSL
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1 MLVSRPFRLRTNNLYLSSMAFSDLLIFLCMPDLDLVRLMQYRPNWFGDLCKLFLF
QY 120 GCTVATLHMTALSVERVLAICRPLARVLVTRRRVZALTVLWAVALLSAGPLLI
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
61 SCTYATVITIALSVRYFAICFPLRAKVWVTKGRVKLVIFVIVWAVAFCSAGGPIF
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
180 EQDPGISVVVGLNGTARIASSPLSSPLMLSRAPPSPSPGPETAEEAALFSGRE
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : :
121 EHE-----NGT-----DP--W-----DINE-----
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
239 -PAQIGALRVMLWVTTYAFFLPFLCLSLYGLIGRELWSSRRPLRGPAASGRERG
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : :
139 FAVRSGLLTVMVVSISFFFLPVFLVLCVLYSLIGRKLWRRRRRGDVAVVGASLRDQNI
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
298 RVLLVWVLAFLICWLPVHGRIIINTEDS----RMVFSQYFNIVALQLQVLSAS
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : :
199 KMLAVVWFAFLICWLPVHGRIYFSKSFEPGSEIAIQISQVCLNLVSVFLYLSAA
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
355 YNLISKYRAAFAKLLILARKSRPFGHRSRDTAGEVAGDTGGDTVGYTETSAN
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Db 259 YNIMSKYRVAVFLLGFBPFSQKSLTLKDESSR-----AWTESSIN 301
199 SLEIAQISQYCNLVSVFVLSAAINPILYNIMSKKYRVAVFLLGFBPFSQKSLTLK 258
386 TAGEVAGDTGGDTVGYTETSAN 407
259 ESSR-----AWTESSIN 270
Search completed: January 1, 2004, 06:48:27
Job time : 19.5539 secs

Db 259 YNIMSKYRVAVFLLGFBPFSQKSLTLKDESSR-----AWTESSIN 301
RESULT 15
US-09-077-675A-12
; Sequence 12, Application US/09077675A
; Patent No. 6242199
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA: US/09/077, 675A
; APPLICATION NUMBER: US/09/077, 675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 271 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-077-675A-12

Query Match 31.3%; Score 674.5; DB 3; Length 271;
Best Local Similarity 45.0%; Pred. No. 2e-50;
Matches 145; Conservative 44; Mismatches 76; Indels 57; Gaps 7;
QY 91 LPFDLVRLWRSPWVFGPLLCRLSLVYGGCTVATLLHMTALSVERVLAICRPLRARVLV 150
DB 1 MPLDLVRLWQYRPNWFGDLCKLFQFVSECTATVLTITALSVERVFAICFPLRAKVVV 60
QY 151 TRRRVRLAIVLWAVALLSAGPFLFLVGVQDPGISVVPGLNGTARIASSPLASSPPLWL 210
DB 61 TKGRVKLVIFVIMAVAFCSAGPIFVLVGEHE-----NGT-----DP--W- 98
QY 211 SRAPPPSPGPGPETAANAALFSRECRPS--PAQLGALRVMLWVTYATFFLPFLCLSLYG 268
DB 99 -----DTNECRPTFAVRSGLLTVMWVSSIEFFLPVFLVLYS 138
QY 269 LIGRELWSSRRPLRGPAAASGRCHROTVRVLLVWVLAFLICWLPFHVGRILIIINTEDS- 327
DB 139 LIGRKLWRRRRGDVAVGASLRDQNHQTVKQVLAIVVFAFILCWLPFHVGRYLFPSKFEFG 198
QY 328 --RMVYFSQYFNIVALQLFVLSASINPILYNLISKYRAAAAFKLLLARKSRPRGFHRSRD 385

THIS PAGE BLANK (USPTO)

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	2155	100.0	412	12	US-10-318-661-28	Sequence 28, Appli
2	2155	100.0	412	12	US-10-206-677-2	Sequence 2, Appli
3	2155	100.0	412	15	US-10-225-567A-473	Sequence 473, Appl
4	2155	100.0	412	15	US-10-290-078-15	Sequence 15, Appl
5	2149	99.7	412	12	US-10-417-820A-130	Sequence 130, Appl
6	2149	99.7	412	12	US-10-417-820A-150	Sequence 150, Appl
7	863.5	40.1	366	12	US-10-303-204A-13	Sequence 13, Appl
8	863.5	40.1	366	15	US-10-251-385-88	Sequence 88, Appl
9	863.4	40.0	353	12	US-10-303-204A-3	Sequence 3, Appli
10	861.5	40.0	364	12	US-10-303-204A-16	Sequence 16, Appli
11	860	39.9	361	12	US-10-303-204A-8	Sequence 8, Appli
12	857.5	39.8	362	15	US-10-251-385-210	Sequence 210, Appl
13	775.5	36.0	302	12	US-10-303-204A-2	Sequence 2, Appli
14	774.5	35.9	302	12	US-10-303-204A-7	Sequence 7, Appli
15	774.5	31.3	271	12	US-10-303-204A-12	Sequence 12, Appl

QY 121 CTYATLLHMTALSVERYLAICRPLRARVLVTRRRVRALIAVLWAVALLSAGPFLVLGVGE 180
DB 121 CTYATLLHMTALSVERYLAICRPLRARVLVTRRRVRALIAVLWAVALLSAGPFLVLGVGE 180
QY 181 QDPGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPGTAAALFSRECRPSA 240
DB 181 QDPGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPGTAAALFSRECRPSA 240
QY 241 QLGALRWMLWVTYAYFFLPCLISILYGLIGRELWSSRRPLRGPAASGRERGHRTQVRVL 300
DB 241 QLGALRWMLWVTYAYFFLPCLISILYGLIGRELWSSRRPLRGPAASGRERGHRTQVRVL 300
QY 301 LVVLAFLIICWLPFHVGRIIYINTEDSRMMYFSQYFNIVALQLFYLSASINPILYNLSK 360
DB 301 LVVLAFLIICWLPFHVGRIIYINTEDSRMMYFSQYFNIVALQLFYLSASINPILYNLSK 360
QY 361 KYRAAFAFKLLARKSRPRGFHRSRDTAGVAGDTGGDTVGYTETSANVKTGM 412
DB 361 KYRAAFAFKLLARKSRPRGFHRSRDTAGVAGDTGGDTVGYTETSANVKTGM 412

RESULT 2

US-10-206-677-2
; Sequence 2, Application US/10206677
; Publication No. US20030186336A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences, Inc.
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burner, Glenna C.
; APPLICANT: Roush, Christine L.
; APPLICANT: Kulander, Bruce G.
; TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC COMPOSITIONS AND METHODS RELATED TO
; FILE REFERENCE: 1920-1-8
; CURRENT APPLICATION NUMBER: US/10/206,677
; CURRENT FILING DATE: 2002-11-06
; PRIOR APPLICATION NUMBER: 60/250,251
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 60/250,452
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-206-677-2

Query Match 100.0%; Score 2155; DB 12; Length 412;
Best Local Similarity 100.0%; Pred. No. 5.4e-167;
Matches 412; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGSPWNGSDGPEGAREPPWPALPPCDERRCSPPLGALVPVTAVCICLVFVGVSGNVTV 60
DB 1 MGSPWNGSDGPEGAREPPWPALPPCDERRCSPPLGALVPVTAVCICLVFVGVSGNVTV 60
QY 61 MLIGRYDRMTTNLYLGSMAVSDLLILGLPDLVLRWSRPWVFGPILCRLSLVXGEG 120
DB 61 MLIGRYDRMTTNLYLGSMAVSDLLILGLPDLVLRWSRPWVFGPILCRLSLVXGEG 120
QY 121 CTYATLLHMTALSVERYLAICRPLRARVLVTRRRVRALIAVLWAVALLSAGPFLVLGVGE 180
DB 121 CTYATLLHMTALSVERYLAICRPLRARVLVTRRRVRALIAVLWAVALLSAGPFLVLGVGE 180
QY 181 QDPGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPGTAAALFSRECRPSA 240
DB 181 QDPGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPGTAAALFSRECRPSA 240
QY 241 QLGALRWMLWVTYAYFFLPCLISILYGLIGRELWSSRRPLRGPAASGRERGHRTQVRVL 300
DB 241 QLGALRWMLWVTYAYFFLPCLISILYGLIGRELWSSRRPLRGPAASGRERGHRTQVRVL 300
QY 301 LVVLAFLIICWLPFHVGRIIYINTEDSRMMYFSQYFNIVALQLFYLSASINPILYNLSK 360
DB 301 LVVLAFLIICWLPFHVGRIIYINTEDSRMMYFSQYFNIVALQLFYLSASINPILYNLSK 360
QY 361 KYRAAFAFKLLARKSRPRGFHRSRDTAGVAGDTGGDTVGYTETSANVKTGM 412
DB 361 KYRAAFAFKLLARKSRPRGFHRSRDTAGVAGDTGGDTVGYTETSANVKTGM 412

DB 301 LVVLAFLIICWLPFHVGRIIYINTEDSRMMYFSQYFNIVALQLFYLSASINPILYNLSK 360
QY 361 KYRAAFAFKLLARKSRPRGFHRSRDTAGVAGDTGGDTVGYTETSANVKTGM 412
DB 361 KYRAAFAFKLLARKSRPRGFHRSRDTAGVAGDTGGDTVGYTETSANVKTGM 412

RESULT 3

US-10-225-567A-473
; Sequence 473, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burner, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 473
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-225-567A-473

Query Match 100.0%; Score 2155; DB 15; Length 412;
Best Local Similarity 100.0%; Pred. No. 5.4e-167;
Matches 412; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGSPWNGSDGPEGAREPPWPALPPCDERRCSPPLGALVPVTAVCICLVFVGVSGNVTV 60
DB 1 MGSPWNGSDGPEGAREPPWPALPPCDERRCSPPLGALVPVTAVCICLVFVGVSGNVTV 60
QY 61 MLIGRYDRMTTNLYLGSMAVSDLLILGLPDLVLRWSRPWVFGPILCRLSLVXGEG 120
DB 61 MLIGRYDRMTTNLYLGSMAVSDLLILGLPDLVLRWSRPWVFGPILCRLSLVXGEG 120
QY 121 CTYATLLHMTALSVERYLAICRPLRARVLVTRRRVRALIAVLWAVALLSAGPFLVLGVGE 180
DB 121 CTYATLLHMTALSVERYLAICRPLRARVLVTRRRVRALIAVLWAVALLSAGPFLVLGVGE 180
QY 181 QDPGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPGTAAALFSRECRPSA 240
DB 181 QDPGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPGTAAALFSRECRPSA 240
QY 241 QLGALRWMLWVTYAYFFLPCLISILYGLIGRELWSSRRPLRGPAASGRERGHRTQVRVL 300
DB 241 QLGALRWMLWVTYAYFFLPCLISILYGLIGRELWSSRRPLRGPAASGRERGHRTQVRVL 300
QY 301 LVVLAFLIICWLPFHVGRIIYINTEDSRMMYFSQYFNIVALQLFYLSASINPILYNLSK 360
DB 301 LVVLAFLIICWLPFHVGRIIYINTEDSRMMYFSQYFNIVALQLFYLSASINPILYNLSK 360
QY 361 KYRAAFAFKLLARKSRPRGFHRSRDTAGVAGDTGGDTVGYTETSANVKTGM 412
DB 361 KYRAAFAFKLLARKSRPRGFHRSRDTAGVAGDTGGDTVGYTETSANVKTGM 412

RESULT 4

US-10-290-078-15
; Sequence 15, Application US/10290078
; Publication No. US20030124596A1
; GENERAL INFORMATION:
; APPLICANT: Carroll, Joseph A.
; TITLE OF INVENTION: Methods and Compositions for Treating
; FILE REFERENCE: Hematological Disorders Using 232, 2059, 10630, 12848, 13875,
; TITLE OF INVENTION: 14395, 14618, 17692 or 58874


```
; FILE REFERENCE: MPI2001-288P1 (M)
; CURRENT APPLICATION NUMBER: US/10/290,078
; CURRENT FILING DATE: 2002-11-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-290-078-15

Query Match      100.0%; Score 2155; DB 15; Length 412;
Best Local Similarity 100.0%; Pred. No. 5.4e-167;
Matches 412; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGSPWNGSDGPEGAREPPWALPPCDERRCSPFPLGALVPVTAACLCFLVGVSGNVTV 60
Db 1 MGSPWNGSDGPEGAREPPWALPPCDERRCSPFPLGALVPVTAACLCFLVGVSGNVTV 60

QY 61 MLIGRYDRMTTNTNLVLSMAVSDLLILLGLPDLVRLWRSRPWFPGPLLCRLSLYVGE 120
Db 61 MLIGRYDRMTTNTNLVLSMAVSDLLILLGLPDLVRLWRSRPWFPGPLLCRLSLYVGE 120

QY 121 CTYATLLHMTALSVERYLAICRPLRARVLTTRRRVRLIAVLWAVALLSAGPFLVGV 180
Db 121 CTYATLLHMTALSVERYLAICRPLRARVLTTRRRVRLIAVLWAVALLSAGPFLVGV 180

QY 181 QDPGISVFLNGTARIASSPLASSPPLWLSRAPPPSPGPTAAEAALFSRECRPSA 240
Db 181 QDPGISVFLNGTARIASSPLASSPPLWLSRAPPPSPGPTAAEAALFSRECRPSA 240

QY 241 QLGALRVMLWTTAYFPFLPCLSLIYGLIGRELWSSRRPLRGPAAASGRGRHQTQV 300
Db 241 QLGALRVMLWTTAYFPFLPCLSLIYGLIGRELWSSRRPLRGPAAASGRGRHQTQV 300

QY 301 LVVLAFLIICWLPFHVGRIIYINTEDSRMMYFSQYFNIVALQFLYLSASINPILYNLISK 360
Db 301 LVVLAFLIICWLPFHVGRIIYINTEDSRMMYFSQYFNIVALQFLYLSASINPILYNLISK 360

QY 361 KYRAAFKLLARKSRPRGFRHSRDTAGEVAGDTGGDTVGYTETSANVKTMG 412
Db 361 KYRAAFKLLARKSRPRGFRHSRDTAGEVAGDTGGDTVGYTETSANVKTMG 412

RESULT 5
US-10-417-820A-130
; Sequence 130, Application US/10417820A
; Publication No. US20030229216A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Ruoping
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lowitz, Kevin
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Behan, Dominic P.
; TITLE OF INVENTION: Constitutively Activated Human G Protein Coupled
; FILE REFERENCE: 7.US28.CON
; CURRENT APPLICATION NUMBER: US/10/417,820A
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: 09/416,760
; PRIOR FILING DATE: 1999-10-12
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: 60/110,060
; PRIOR FILING DATE: 1998-11-27
; PRIOR APPLICATION NUMBER: 60/120,416
; PRIOR FILING DATE: 1999-02-16
; PRIOR APPLICATION NUMBER: 60/121,852
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 60/109,213
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/123,944
; PRIOR FILING DATE: 1999-03-12

; FILE REFERENCE: MPI2001-288P1 (M)
; CURRENT APPLICATION NUMBER: US/10/290,078
; CURRENT FILING DATE: 2002-11-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-290-078-15

Query Match      100.0%; Score 2155; DB 15; Length 412;
Best Local Similarity 100.0%; Pred. No. 5.4e-167;
Matches 412; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGSPWNGSDGPEGAREPPWALPPCDERRCSPFPLGALVPVTAACLCFLVGVSGNVTV 60
Db 1 MGSPWNGSDGPEGAREPPWALPPCDERRCSPFPLGALVPVTAACLCFLVGVSGNVTV 60

QY 61 MLIGRYDRMTTNTNLVLSMAVSDLLILLGLPDLVRLWRSRPWFPGPLLCRLSLYVGE 120
Db 61 MLIGRYDRMTTNTNLVLSMAVSDLLILLGLPDLVRLWRSRPWFPGPLLCRLSLYVGE 120

QY 121 CTYATLLHMTALSVERYLAICRPLRARVLTTRRRVRLIAVLWAVALLSAGPFLVGV 180
Db 121 CTYATLLHMTALSVERYLAICRPLRARVLTTRRRVRLIAVLWAVALLSAGPFLVGV 180

QY 181 QDPGISVFLNGTARIASSPLASSPPLWLSRAPPPSPGPTAAEAALFSRECRPSA 240
Db 181 QDPGISVFLNGTARIASSPLASSPPLWLSRAPPPSPGPTAAEAALFSRECRPSA 240

QY 241 QLGALRVMLWTTAYFPFLPCLSLIYGLIGRELWSSRRPLRGPAAASGRGRHQTQV 300
Db 241 QLGALRVMLWTTAYFPFLPCLSLIYGLIGRELWSSRRPLRGPAAASGRGRHQTQV 300

QY 301 LVVLAFLIICWLPFHVGRIIYINTEDSRMMYFSQYFNIVALQFLYLSASINPILYNLISK 360
Db 301 LVVLAFLIICWLPFHVGRIIYINTEDSRMMYFSQYFNIVALQFLYLSASINPILYNLISK 360

QY 361 KYRAAFKLLARKSRPRGFRHSRDTAGEVAGDTGGDTVGYTETSANVKTMG 412
Db 361 KYRAAFKLLARKSRPRGFRHSRDTAGEVAGDTGGDTVGYTETSANVKTMG 412

RESULT 6
US-10-417-820A-150
; Sequence 150, Application US/10417820A
; Publication No. US20030229216A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Ruoping
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lowitz, Kevin
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Behan, Dominic P.
; TITLE OF INVENTION: Constitutively Activated Human G Protein Coupled
; FILE REFERENCE: 7.US28.CON
; CURRENT APPLICATION NUMBER: US/10/417,820A
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: 09/416,760
; PRIOR FILING DATE: 1999-10-12
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: 60/110,060
; PRIOR FILING DATE: 1998-11-27
; PRIOR APPLICATION NUMBER: 60/120,416
; PRIOR FILING DATE: 1999-02-16
; PRIOR APPLICATION NUMBER: 60/121,852
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 60/123,944
; PRIOR FILING DATE: 1999-03-12
```

```
; PRIOR APPLICATION NUMBER: 60/109,213
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/123,944
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/123,945
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/123,948
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/123,951
; PRIOR FILING DATE: 1999-03-12
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 150
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-417-820A-150

Query Match
Best Local Similarity 99.8%; Score 2149; DB 12; Length 412;
Matches 411; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGSPWNGSDGEGAREPPWPPALPCDERRCSPPLGALVPTAVCLFVGVSGNVTV 60
Db 1 MGSPWNGSDGEGAREPPWPPALPCDERRCSPPLGALVPTAVCLFVGVSGNVTV 60
QY 61 MLIGRYDRMTTNNLYLGSMAVSDLLILGLPDLRLWRSRPWVFGPCLLCRLSLYVGE 120
Db 61 MLIGRYDRMTTNNLYLGSMAVSDLLILGLPDLRLWRSRPWVFGPCLLCRLSLYVGE 120
QY 121 CTYATLLHMTALSVERYLAICRPLRARLVTRRRVRALIAVLWAVALLSAGPFLVGVGE 180
Db 121 CTYATLLHMTALSVERYLAICRPLRARLVTRRRVRALIAVLWAVALLSAGPFLVGVGE 180
QY 181 QDPGISVPGCLNGTARIASSPLASSPPLWLSRAPPPSPGPGTAAALFSRECRPSA 240
Db 181 QDPGISVPGCLNGTARIASSPLASSPPLWLSRAPPPSPGPGTAAALFSRECRPSA 240
QY 241 QLGALRVMLVTTAYFPLPCILSILYGLIGRLWSSRRPLRGPAAASGRGHRQTVRL 300
Db 241 QLGALRVMLVTTAYFPLPCILSILYGLIGRLWSSRRPLRGPAAASGRGHRQTVRL 300
QY 301 LVVLAFLIICWLPFHVGRIIYINTEDSRMMYFQYFNIVALQFLYLSASINPILYNLSK 360
Db 301 LVVLAFLIICWLPFHVGRIIYINTEDSRMMYFQYFNIVALQFLYLSASINPILYNLSK 360
QY 361 KYRAAFAKLLARKSRPRGHRSDTAGEVAGDTGGDTVGYTETSANVKTMG 412
Db 361 KYRAAFAKLLARKSRPRGHRSDTAGEVAGDTGGDTVGYTETSANVKTMG 412
```

RESULT 7

```
US-10-303-204A-13
; Sequence 13, Application US/10303204A
; Publication No. US20030166144A1
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus H. T.
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR
; FILE REFERENCE: 19589PCA
; CURRENT APPLICATION NUMBER: US/10/303,204A
; CURRENT FILING DATE: 2002-11-25
; PRIOR APPLICATION NUMBER: 09/077,674
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: PCT/US96/19445
; PRIOR FILING DATE: 1996-12-10
```

```
; PRIOR APPLICATION NUMBER: 60/018,962
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: 60/008,582
; PRIOR FILING DATE: 1995-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-303-204A-13
```

```
Query Match
Best Local Similarity 40.1%; Score 863.5; DB 12; Length 366;
Matches 186; Conservative 59; Mismatches 106; Indels 65; Gaps 9;

QY 5 WNGSDGEGA-----REPPWALPPCD---ERRCSPPLGALVPTAVCLCFLFVGVSGN 56
Db 2 WNPATSEEPFNLTADLDWDASPGNDSLGDELLQLFPAPLLAGVATATCVAFVVGAGN 61
QY 57 VVTVMILIGRYDRMTTNNLYLGSMAVSDLLILGLPDLRLWRSRPWVFGPCLLCRLSLY 116
Db 62 LUTMLVWSRFRRLTNNLYLSSMAFSDLLIFLCMPDLRLWQYRPMWPGDLLCKLQF 121
QY 117 VGEGETYATLLHMTALSVERYLAICRPLRARLVTRRRVRALIAVLWAVALLSAGPFLFL 176
Db 122 VSECTYATVLTALSVERYFAICFPLRAKVVTGKRVKLVIFVIWAVAFCSAGPFLVL 181
QY 177 VGVQDPGISVPGCLNGTARIASSPLASSPPLWLSRAPPPSPGPGTAAALFSRECR 236
Db 182 VGVGEHE-----NGT-----DP--W-----DTNECR 199
QY 237 PS--PAQLGALRVMLVTTAYFPLPCILSILYGLIGRLWSSRRPLRGPAAASGRGHR 294
Db 200 PTEFAVRSGLLTVMWVSSIFFPLPVFCITVLYSLIGRLWRRRRGDVVGASLRDQNHK 259
QY 295 QTVRVLLVVLVAFIICWLPFHVGRIIYINTEDS---RMMYFQYFNIVALQFLYLSASIN 351
Db 260 QIVKMLAVVFAFILLWLPFHVGRIYLFKSPGSGLSIAQISQYCNLVSVFLYLSA 319
QY 352 PILYNIISKYRAAFAKLLARKSRPRGHRSDTAGEVAGDTGGDTVGYTETSAN 407
Db 320 PILYNIIMSKYRAVAVFLLGFPFQSKLSTLKDESSR-----AWTESSIN 365
```

RESULT 8

```
US-10-251-385-88
; Sequence 88, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
; TITLE OF INVENTION: Protein-Coupled
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/10/251,385
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US/09/170,496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 88
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-251-385-88
```

```
Query Match
Best Local Similarity 40.1%; Score 863.5; DB 15; Length 366;
Matches 186; Conservative 59; Mismatches 106; Indels 65; Gaps 9;
```

```
QY 5 WNGSDGPEGA-----REPPALPPCD-----ERRCSPPFPGALVPVTVAVCLCLFVVGUSGN 56
DB 2 WNATPSEEPFNLTTLADLDWDASPGNDSGLDELQLFPAPLLAGVATATCVALLVVGVIAGN 61
QY 57 VVTVMIGRYDRMTTNNLYLGSMAVSDLLILGLPFDLYLRWSRPWVFGPLLCRLSLY 116
DB 62 LUTMLVVSFRRLTNNLYLSMAFSDDLIFLCMLDLVRLWQYRPWFGDLCCKLQF 121
QY 117 VGEGETYATLLHMTALSVRYLAICRPLRARVLTTRRRVRLIAVLWAVALLSAGPFLF 176
DB 122 VSESCTYATVLTITALSVERYPFAICPLRAKVVVTKGRVKLVIFVIWAVAFCSAGPIFVL 181
QY 177 VGVQDPGIVVPGVGLNGTARIASSPLASSPPLWLRAPPSPGPGPETABAAALFSRECR 236
DB 182 VGVGHE-----NGT-----DP--W-----DTNECR 199
QY 237 PS--PAQCALRVMLWVTTAYPFLPCLCLSLYGLIGRELSSRRPLRGPAASGRERGRH 294
DB 200 PTEFAVRSGLLTVMWVSSIFPFLPVFCUTVLSLIGRLWRRRRGDVVGASLRQDNHK 259
QY 295 QTVRVLLVVLAFIICWLPFHVGRIIYINTEDS---RMMYFSQYFNIVALQLFYLSASIN 351
DB 260 QTVKMLAVVVFILCMLPFHVGRLYFSKFEPSGLSLEIAISQYCNLVSFVLFYLSAAN 319
QY 352 PILYNLSIKKYRAAARFKLLARKSRPRGHRSDTAGVAGDTGGDTVGYTETSAN 407
DB 320 PILYNMSKKYRVAVFRLGLFEPFSQKSLTKDESSR-----AWTESSIN 365

RESULT 9
US-10-303-204A-3
; Sequence 3, Application US/10303204A
; Publication No. US20030166144A1
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feigmer, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus H. T.
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR
; TITLE OF INVENTION: FAMILY
; FILE REFERENCE: 19589PCA
; CURRENT APPLICATION NUMBER: US/10/303,204A
; CURRENT FILING DATE: 2002-11-25
; PRIOR FILING DATE: 2002-11-25
; PRIOR APPLICATION NUMBER: 09/077,674
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: PCT/US96/19445
; PRIOR FILING DATE: 1996-12-10
; PRIOR APPLICATION NUMBER: 60/018,962
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: 60/008,582
; PRIOR FILING DATE: 1995-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 353
; TYPE: PRT
; ORGANISM: sus scrofa
US-10-303-204A-3

Query Match 40.0%; Score 863; DB 12; Length 353;
Best Local Similarity 46.5%; Pred. No. 4.1e-62;
Matches 186; Conservative 53; Mismatches 95; Indels 66; Gaps 9;

QY 19 WPALPPCD---ERRCSPPFPGALVPVTVAVCLCLFVVGUSGNVVTMLIGRYDRMTTNNL 75
DB 8 WDAPPENDSLVEBLLPFLPFPVLLAGVATATCVALLVVGVIAGNLTMLVVSFRFWRMTTNNL 67
QY 76 YLGSMAVSDLLILGLPFDLYLRWSRPWVFGPLLCRLSLYVGEGETYATLLHMTALSVE 135
DB 68 YLSSMAFSDDLIFLCMLDLFRLWQYRPWNLGNLCKLCKLQFVSESCTYATVLTITALSVE 127
```

```
QY 136 RYLAICRPLRARVLTTRRRVRLIAVLWAVALLSAGPFLFVGVQDPGIVVVGUSGN 195
DB 128 RYFAICPPLRAKVVVTKGRVKLVIFVIWAVAFCSAGPIFVLVGVGHE-----NGT- 177
QY 196 RIASSPLASSPPLWLRAPPSPGPGPETABAAALFSRECRPS--PAQCALRVMLWVTT 253
DB 178 -----DPRD-----TNECRATEFAVRSGLLTVMWVSS 205
QY 254 AYFFLPFLCLSLYGLIGRELSSRRPLRGPAASG---RERGHROTVRVLLVVLAFIIC 310
DB 206 VFFPFLPVCLTVLSLIGRLW---RRKGEAAVSSLRDQNHKQTVKMLAVVVPFAPILC 262
QY 311 WLPFHVGRIIY---INTEDSRMMYFSQYFNIVALQLFYLSASINPILYNLSIKKYRAAAR 367
DB 263 WLPFHVGRLYFSKLEPSGVEIAQISQYCNLVSFVLFYLSAANPILYNLSIKKYRVAVF 322
QY 368 KULLARKSRPRGHRSDTAGVAGDTGGDTVGYTETSAN 407
DB 323 KULGPEPFSQKSLTKDESSR-----AWTESSIN 352

RESULT 10
US-10-303-204A-16
; Sequence 16, Application US/10303204A
; Publication No. US20030166144A1
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris P.
; APPLICANT: Feigmer, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus H. T.
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR
; TITLE OF INVENTION: FAMILY
; FILE REFERENCE: 19589PCA
; CURRENT APPLICATION NUMBER: US/10/303,204A
; CURRENT FILING DATE: 2002-11-25
; PRIOR FILING DATE: 09/077,674
; PRIOR APPLICATION NUMBER: PCT/US96/19445
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/018,962
; PRIOR FILING DATE: 1996-12-10
; PRIOR APPLICATION NUMBER: 60/008,582
; PRIOR FILING DATE: 1995-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 364
; TYPE: PRT
; ORGANISM: rattus norvegicus
US-10-303-204A-16

Query Match 40.0%; Score 861.5; DB 12; Length 364;
Best Local Similarity 48.3%; Pred. No. 5.7e-62;
Matches 185; Conservative 52; Mismatches 81; Indels 65; Gaps 11;

QY 5 WNGSDGPEGARPP-----WPALPPCD---ERRCSPPFPGALVPVTVAVCLCLFVVGUSG 55
DB 2 WNAT--PSEPEPNNVTLDDWDASPGNDSLPDBELLPLFPAPLLAGVATATCVALLVVGISG 59
QY 56 NVVTVMIGRYDRMTTNNLYLGSMAVSDLLILGLPFDLYLRWSRPWVFGPLLCRLSL 115
DB 60 NLLTMLVVSFRRLTNNLYLSMAFSDDLIFLCMLDLVRLWQYRPWNLGNLCKLQF 119
QY 116 YVGEGETYATLLHMTALSVERYPFAICRPLRARVLTTRRRVRLIAVLWAVALLSAGPFLF 175
DB 120 FVSESCTYATVLTITALSVERYPFAICPLRAKVVVTKGRVKLVIFVIWAVAFCSAGPIFV 179
QY 176 LVGVQDPGIVVVGUSGNVVTMLIGRYDRMTTNNLYLGSMAVSDLLILGLPFDLYLRWSRP 235
```

Db 180 LVGVEHE-----NGT-----DPRD-----TNEC 197
 QY 236 RPS--PAQLGALRVMLVWVTTAYFPFLPCLSLILYGLIGRELWSSRPLRGPAAG--RE 290
 Db 198 RATEFAVRSGLLTVWVWSSVFFLPVFCVTVLSLIGRKLW--RR--RGDAAVGASLRD 253
 QY 291 RGHRTQVRVLLVVLVLAFLIICWLPFHVGRIRIYINTEDS---RMVFSQYFNIVALQLFYLS 347
 Db 254 QNHKQIVKMLAVVVFALICWLPFHVGRYLFKSPFPGSLIAQISOYCNLVSVFLFYLS 313
 QY 348 ASINPILYNISKYRAAAFKLL 370
 Db 314 AAINPILYNIMSKKYRVAVFKLL 336

RESULT 11

US-10-303-204A-8
 ; Sequence 8, Application US/10303204A
 ; Publication No. US20030166144A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Arena, Joseph P.
 ; APPLICANT: Cully, Doris F.
 ; APPLICANT: Feighner, Scott D.
 ; APPLICANT: Howard, Andrew D.
 ; APPLICANT: Liberator, Paul A.
 ; APPLICANT: Schaeffer, James M.
 ; APPLICANT: Van Der Ploeg, Leonardus H. T.
 ; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR
 ; FILE REFERENCE: 19589PCA
 ; CURRENT FILING DATE: 2002-11-25
 ; PRIOR FILING DATE: 2002-11-25
 ; PRIOR FILING DATE: 1998-06-03
 ; PRIOR FILING DATE: 1996-12-10
 ; PRIOR FILING DATE: 1996-06-06
 ; PRIOR FILING DATE: 1996-06-06
 ; PRIOR FILING DATE: 1995-12-13
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 8
 ; LENGTH: 361
 ; TYPE: PRT
 ; ORGANISM: homo sapiens
 US-10-303-204A-8

Query Match 39.9%; Score 860; DB 12; Length 361;
 Best Local Similarity 46.1%; Pred. No. 7.4e-62;
 Matches 183; Conservative 57; Mismatches 97; Indels 60; Gaps 8;
 QY 19 WPALPPCD---ERRCSPFPLGALVPVTVAVCLCLFVVGSGNVVTVMLIGRYRDMRTTNL 75
 Db 16 WDASPGNDSLGDELLQFPAPLAGVATCVALFVVGAGNLLTMLVSRFRELRTTNL 75
 QY 76 YLGSMVSDLLILGLPFDLYLWRSPWPVFGPLLCRLSYVCEGCTYATLLHMTALSVE 135
 Db 76 YLGSMVSDLLILGLPFDLYLWRSPWPVFGPLLCRLSYVCEGCTYATLLHMTALSVE 135
 QY 136 RYLAICRPLRARLVTRRRVRALIAVLWALLSAGPFLVGVGQDPGISVVPGLNGTA 195
 Db 136 RYFAICFPLRAKVTVTKRVLKLVFIWAVAFCSAGPIFVLVGVGHE-----NGT- 185
 QY 196 RIASSPLASSPPLWLSRAPPPSPGPETAEEAALFSRECRPS--PAQLGALRVMLVWTT 253
 Db 186 -----DP--W-----DTNECRPTEFAVRSGLLTVWVWVSS 213
 QY 254 AVEFFPLCLSLILYGLIGRELWSSRPLRGPAAGRGHRTQVRVLLVVLVLAFLICWLP 313
 Db 214 IFPFLPVFCVTVLSLIGRKLWRRRGDAVVGASLRDQNHKQIVKMLAVVVFALICWLP 273
 QY 314 FHVGRIRIYINTEDS---RMVFSQYFNIVALQLFYLSASINPILYNISKYRAAAFKLL 370

Db 274 FHVGRYLFKSPFPGSLIAQISOYCNLVSVFLFYLSAALNPILYNIMSKKYRVAVFKLL 333
 QY 371 LARKSPRGFHRSDRTAGEVAGDTGGDTVGYTETSAN 407
 Db 334 GPEFPFSQRLKSTLKDESSR-----AWTESSIN 360
 RESULT 12
 US-10-251-385-210
 ; Sequence 210, Application US/10251385
 ; Publication No. US20030105292A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Behan, Dominic P.
 ; APPLICANT: Chalmers, Derek T.
 ; APPLICANT: Liaw, Chen W.
 ; TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
 ; TITLE OF INVENTION: Protein-Coupled
 ; TITLE OF INVENTION: Receptors
 ; FILE REFERENCE: AREN-0040
 ; CURRENT APPLICATION NUMBER: US/10/251,385
 ; CURRENT FILING DATE: 2002-09-20
 ; PRIOR APPLICATION NUMBER: US/09/170,496
 ; PRIOR FILING DATE: 1998-10-13
 ; NUMBER OF SEQ ID NOS: 294
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 210
 ; LENGTH: 366
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-251-385-210

Query Match 39.8%; Score 857.5; DB 15; Length 366;
 Best Local Similarity 44.5%; Pred. No. 1.2e-61;
 Matches 185; Conservative 59; Mismatches 107; Indels 65; Gaps 9;
 QY 5 WNGSDGPEGA---REPPWALPPCD---ERRCSPFPLGALVPVTVAVCLCLFVVGSGN 56
 Db 2 WNA TPSEEPGFNLTLADLWDASPGNDSLGDELLQFPAPLAGVATCVALFVVGAGN 61
 QY 57 VVTVMILIGRYRDMRTTNLYLGSMVSDLLILGLPFDLYLWRSPWPVFGPPLLCRLSLY 116
 Db 62 LTLMLVSRFRELRTTNLYLGSMVSDLLILGLPFDLYLWRSPWPVFGPPLLCRLSLY 121
 QY 117 VGEGCTYATLLHMTALSVERYLAICRPLRARLVTRRRVRALIAVLWALLSAGPFL 176
 Db 122 VSECTYATVLTITALSVERYFAICFPLRAKVTVTKRVLKLVFIWAVAFCSAGPIFVL 181
 QY 177 VGVGQDPGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPETAEEAALFSRECR 236
 Db 182 VGVGHE-----NGT-----DP--W-----DTNECR 199
 QY 237 PS--PAQLGALRVMLVWTTAYFPFLPCLSLILYGLIGRELWSSRPLRGPAAGRGHR 294
 Db 200 PTEFAVRSGLLTVWVWSSVFFLPVFCVTVLSLIGRKLWRRRGDAVVGASLRDQNHK 259
 QY 295 QTVRVLLVVLVLAFLIICWLPFHVGRIRIYINTEDS---RMVFSQYFNIVALQLFYLSIN 351
 Db 260 QTKMLAVVVFALICWLPFHVGRYLFKSPFPGSLIAQISOYCNLVSVFLFYLSAAN 319
 QY 352 PILYNLISKYRAAAFKLLLARKSRPRGHRSDRTAGEVAGDTGGDTVGYTETSAN 407
 Db 320 PILYNIMSKKYRVAVFKLLGPEFPFSQRLKSTLKDESSR-----AWTESSIN 365

RESULT 13

US-10-303-204A-2
 ; Sequence 2, Application US/10303204A
 ; Publication No. US20030166144A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Arena, Joseph P.
 ; APPLICANT: Cully, Doris F.
 ; APPLICANT: Feighner, Scott D.

```
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus H. T.
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR
; TITLE OF INVENTION: FAMILY
; FILE REFERENCE: 19589PCA
; CURRENT APPLICATION NUMBER: US/10/303,204A
; CURRENT FILING DATE: 2002-11-25
; PRIOR APPLICATION NUMBER: 09/077,674
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: PCT/US96/19445
; PRIOR FILING DATE: 1996-12-10
; PRIOR APPLICATION NUMBER: 60/018,962
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: 60/008,582
; PRIOR FILING DATE: 1995-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 302
; TYPE: PRT
; ORGANISM: eus scrofa
US-10-303-204A-2

Query Match      36.0%; Score 775.5; DB 12; Length 302;
Best Local Similarity 46.6%; Pred. No. 4.4e-55;
Matches 166; Conservative 48; Mismatches 79; Indels 63; Gaps 8;

QY 60 VMLIGRYDMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSPWVFGPLLCRLSLYVGE 119
Db 1 MLVSRFRRLRTTNNLYLSSMAFSDLLIFLCMPDLVRLQYRPMNGLNLLCKLFQVSE 60
QY 120 GCTYATLLHMTALSVERYLAIICPLARVLVTRRRVRLALIVLWAVALLSAGPFLVLGV 179
Db 61 SCTYATVLTITALSVERYFAICPLRAKVVTGKRVKLVIVFWAVAFCSAGPFLVLGV 120
QY 180 EODPGISVVGNGTARIASSPLASPPPLWLSRAPPPSPGPETAEEAALFSRECRPS- 238
Db 121 EHD-----NGT-----DP--W-----DTECRPTE 138
QY 239 -PAQGLALRVMLWVTYAYFPFLPCLSLYGLIGRELWSSRRPLRGPAASG-----RRGRHR 294
Db 139 FAVRSGLLTVMWVSSVFFLPVFCULTVLYSLIGRKLW-----RRKRGAAVGSRLRQNHK 195
QY 295 QTVRVLLVVVLAFLICWLPHVGRITIIY---INTEDSRMVFYFOYFNIVALQFVLSASIN 351
Db 196 QTVKMLAVVVFALICWLPHVGRYLFKSLRPGSVVEIAQISQYCNLVSVFLVLSAAIN 255
QY 352 PILYNLISKYRAAFAKLLARKSRPRGHRSDTAGVAGDTGGDTGVGTETSAN 407
Db 256 PILYNIMSKYRVAVFKLLGFEPFSQKLSLKDESSR-----AWTESSIN 301

RESULT 14
US-10-303-204A-7
; Sequence 7, Application US/10303204A
; Publication No. US20030166144A1
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus H. T.
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR
; TITLE OF INVENTION: FAMILY
; FILE REFERENCE: 19589PCA
; CURRENT APPLICATION NUMBER: US/10/303,204A
; CURRENT FILING DATE: 2002-11-25
; PRIOR APPLICATION NUMBER: 09/077,674
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: PCT/US96/19445
; PRIOR FILING DATE: 1996-12-10
; PRIOR APPLICATION NUMBER: 60/018,962
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: 60/008,582
; PRIOR FILING DATE: 1995-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 271
; TYPE: PRT

; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus H. T.
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR
; TITLE OF INVENTION: FAMILY
; FILE REFERENCE: 19589PCA
; CURRENT APPLICATION NUMBER: US/10/303,204A
; CURRENT FILING DATE: 2002-11-25
; PRIOR APPLICATION NUMBER: 09/077,674
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: PCT/US96/19445
; PRIOR FILING DATE: 1996-12-10
; PRIOR APPLICATION NUMBER: 60/018,962
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: 60/008,582
; PRIOR FILING DATE: 1995-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 302
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-303-204A-7

Query Match      35.9%; Score 774.5; DB 12; Length 302;
Best Local Similarity 46.5%; Pred. No. 5.3e-55;
Matches 164; Conservative 51; Mismatches 81; Indels 57; Gaps 7;

QY 60 VMLIGRYDMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSPWVFGPLLCRLSLYVGE 119
Db 1 MLVSRFRRLRTTNNLYLSSMAFSDLLIFLCMPDLVRLQYRPMNFGDLCKLFQVSE 60
QY 120 GCTYATLLHMTALSVERYLAIICPLARVLVTRRRVRLALIVLWAVALLSAGPFLVLGV 179
Db 61 SCTYATVLTITALSVERYFAICPLRAKVVTGKRVKLVIVFWAVAFCSAGPFLVLGV 120
QY 180 EODPGISVVGNGTARIASSPLASPPPLWLSRAPPPSPGPETAEEAALFSRECRPS- 238
Db 121 EHD-----NGT-----DP--W-----DTECRPTE 138
QY 239 -PAQGLALRVMLWVTYAYFPFLPCLSLYGLIGRELWSSRRPLRGPAASGRRGRHRQTV 297
Db 139 FAVRSGLLTVMWVSSVFFLPVFCULTVLYSLIGRKLWRRRGDAVVGASLRDQNHQTV 198
QY 298 RVLLVVVLAFLICWLPHVGRITIIYINTEDS---RMVFYFOYFNIVALQFVLSASINPIL 354
Db 199 KMLAVVVFALICWLPHVGRYLFKSLRPGSVVEIAQISQYCNLVSVFLVLSAAINPIL 258
QY 355 YNLISKYRAAFAKLLARKSRPRGHRSDTAGVAGDTGGDTGVGTETSAN 407
Db 259 YNIMSKYRVAVFKLLGFEPFSQKLSLKDESSR-----AWTESSIN 301

RESULT 15
US-10-303-204A-12
; Sequence 12, Application US/10303204A
; Publication No. US20030166144A1
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus H. T.
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR
; TITLE OF INVENTION: FAMILY
; FILE REFERENCE: 19589PCA
; CURRENT APPLICATION NUMBER: US/10/303,204A
; CURRENT FILING DATE: 2002-11-25
; PRIOR APPLICATION NUMBER: 09/077,674
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: PCT/US96/19445
; PRIOR FILING DATE: 1996-12-10
; PRIOR APPLICATION NUMBER: 60/018,962
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: 60/008,582
; PRIOR FILING DATE: 1995-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 271
; TYPE: PRT
```

```

; ORGANISM: homo sapiens
US-10-303-204A-12
Query Match      31.3%; Score 674.5; DB 12; Length 271;
Best Local Similarity 45.0%; Pred.No. 6.3e-47;
Matches 145; Conservative 44; Mismatches 76; Indels 57; Gaps 7;

QY 181 LPFDLYRLWRSRPWVFGPLLCRLSLVYEGECTYATLIHMTALSVERYLAICRPLRARVLV 150
Db 181 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 151 TRRRVRLIALVLAVALLSAGPFLVGVQDQPGISVVPGLNGTARIASSPLASSPLML 210
Db 151 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 TKGRVKLVIFVIWAVAFCSAGPIFVLVGEHE-----NGT-----DP--W- 98
QY 211 SRAPPSPSPGPETAEEAALFSBECRPS--PAQLGALRVMLWTTAYFELPFLCLSLY 268
Db 99 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 269 LIGRELWSSRRPLRGPAASGRGRHQTVRVLVWVLAFTICWLPFHVGRIIYINTEDS- 327
Db 139 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 328 --RMVYFQYFNIVALQLFYLSASINPILYNLISKYRAAFKLLARKSRPRGFRSRD 385
Db 199 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 386 TAGEVAGDTGGDTVGYTETSAN 407
Db 259 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
;-----AWTESSIN 270
```

Search completed: January 1, 2004, 06:53:28
Job time : 152.273 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 1, 2004, 06:20:00 ; Search time 26.3308 Seconds
(without alignments)
1504.757 Million cell updates/sec

Title: US-09-719-485-3
Perfect score: 2155
Sequence: 1 MGSPWNGSDGPEGAREPPWP.....DTGGDTVGYTTSANVKTWG 412

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	501	23.2	424	2 JH0164	neurotensin recept
2	497	23.1	418	2 S29506	neurotensin recept
3	453	21.0	477	2 JC7313	capa receptor (CGI
4	437	20.3	418	2 A88013	protein K10B4.4 [i
5	401	18.6	416	2 S68822	neurotensin recept
6	375	17.4	378	2 T15816	hypothetical prote
7	359	16.7	363	2 I57940	somatostatin recep
8	357.5	16.6	418	2 A46226	somatostatin recep
9	344	16.0	363	2 I57955	somatostatin recep
10	344	16.0	364	2 JN0763	somatostatin recep
11	343.5	15.9	380	2 A55259	kappa opioid recep
12	339	15.7	352	2 JE0296	thyrotropin releas
13	339	15.7	380	2 S36143	kappa opioid recep
14	335	15.5	380	2 JC2338	kappa opioid recep
15	335	15.5	388	2 JN0605	somatostatin recep
16	334.5	15.5	367	2 I49022	kappa opioid recep
17	334.5	15.5	367	2 JC2421	opioid receptor ho
18	334.5	15.5	367	2 I56520	G protein-coupled
19	334.5	15.5	428	2 S30508	probable G protein
20	331.5	15.4	428	2 A44021	somatostatin recep
21	330.5	15.3	370	2 S43087	orphan opioid rece
22	329	15.3	380	2 A48227	kappa opioid recep
23	328	15.2	380	2 JC2434	kappa opioid recep
24	327.5	15.2	384	2 A47249	brain-specific som
25	326.5	15.2	372	2 I38532	delta opioid recep
26	326	15.1	519	2 S17783	tachykinin recept
27	322	14.9	384	2 JC4629	somatostatin recep
28	321.5	14.9	372	2 S34592	delta opioid recep
29	321.5	14.9	398	2 JN0708	thyrotropin-releas

RESULT 1

JH0164

neurotensin receptor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 17-Mar-2000

C:Accession: JH0164

R:Tanaka, K.; Maou, M.; Nakanishi, S.

Neuron 4, 847-854, 1990

A:Title: Structure and functional expression of the cloned rat neurotensin receptor.

A:Reference number: JH0164; PMID:1694443

A:Accession: JH0164

A:Molecule type: mRNA

A:Residues: 1-424 <TAN>

C:Comment: Neurotensin receptor belongs to the family of G protein-coupled receptor. This

cter (neurotransmitter) in the brain and as a hormone) cellular mediator in peripheral tiss

C:Superfamily: vertebrate rhodopsin

C:Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein

F:65-87/Domain: transmembrane #status predicted <TM1>

F:97-121/Domain: transmembrane #status predicted <TM2>

F:144-165/Domain: transmembrane #status predicted <TM3>

F:189-210/Domain: transmembrane #status predicted <TM4>

F:236-260/Domain: transmembrane #status predicted <TM5>

F:309-330/Domain: transmembrane #status predicted <TM6>

F:348-372/Domain: transmembrane #status predicted <TM7>

F:4,38,42/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 23.2%; Score 501; DB 2; Length 424;

Best Local Similarity 32.9%; Pred. No. 2.4e-32;

Matches 125; Conservative 59; Mismatches 110; Indels 86; Gaps 11;

QY	39	VPTAVTCLFVGVGSGNVTVMLIGR---	YRDMRTTNLYLGSMAVSLLI-LLGLPFD 94
DB	65	VLVTATYALFVTVGVGSGNVTVMLIGR---	YRDMRTTNLYLGSMAVSLLI-LLGLPFD 94
QY	95	LVR-LWRSPWVFGPLCLSLVVGCGCTVATLHMTALSVERYLAICRPLRVLVTRR 153	
DB	125	LYNFIWHPWAFGDCAGCYFLRDACTYATALNVASLSVERYLAICHPFKATLMRSR 184	
QY	154	RVRALIAVLWAVALLSAGPFLPLVGVQDPGISVWPLNGTARIASPLASSPPLMSRA 213	
DB	185	RTKKFISATWLASALLAIPMLFTMGILQNRSGDGTGPG----	GLVCTPIVDT----- 231
QY	214	PPSPSPSGPETAAPALFSECRPSPAQIGALRVLWLVVTTAYFFL-PFLCLSLYGLIGR 272	
DB	232	-----ATVKVIVQVNTFMSFLPFLPMLVISILNTVIAN 262	
QY	273	ELWSSRRRLRGPAAAGR-----	ERGHQTVR---VLLVVLAFII 309
DB	263	KLTVM---VHQAAEQQRVCTVGTGNGLEHSTFNWTEPGRVQALRGVLVRAVIAFVV 319	
QY	310	CWLPPFHVGRIIYINTEDSR-----	MMYFSQYFNIVALQLPYLSASINPILYNLISKYRAA 365

:

```
Db 320 CWLPYHVRRLMFCYISDEQWTTFLDFYHYFYMLTNLFYVSSAINPILNLYSANFRQV 379
QY 366 AFKLLLA-----RKSRP 377
Db 380 FLSTLACLCPGWRHRRKRP 399

RESULT 2
S29506
neurotensin receptor - human
C;Species: Homo sapiens (man)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 17-Mar-2000
C;Accession: S29506
R;Vita, N.; Laurent, P.; Lefort, S.; Chalon, P.; Dumont, X.; Kaghad, M.; Gully, D.; le H
FEBS Lett. 317, 139-142, 1993
A;Title: Cloning and expression of a complementary DNA encoding a high affinity human ne
A;Reference number: S29506; MUID:93154505; PMID:8361365
A;Accession: S29506
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-418 <VT>
A;Cross-references: EMBL:X70070; NID:g35020; PIDN:CAA49675.1; PID:g35021
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; transmembrane protein

Query Match 23.1%; Score 497; DB 2; Length 418;
Best Local Similarity 33.1%; Pred. No. 5e-32;
Matches 126; Conservative 62; Mismatches 117; Indels 76; Gaps 13;

QY 39 VPTAVCLCLFVVGSGNVVTVMLIGR---YRDMRTTNNLYLGSMAVSDLL-ILLGLPFD 94
Db 64 VLVTVAVLALFVVGTVGNTVTAFTLARKKSLSQSLQSTVHVLGSLSDLLTLLAMPVE 123
QY 95 LYR-LWRSRPPVGPCLLCRLSLVGGCTVATLLHMTALSVERYLAICRPLRARVLVTRR 153
Db 124 LYNFIWVHHWFAPGDACRGYFLRDACVATALNVAASLVERYLAICHPFKATLMRS 183
QY 154 RVRLATLAVLWALLSAGPFLFVGVEQDPSVVPGLNGTARIASSPLSSPPLWLSRA 213
Db 184 RTKKFSAIWLASALLTVMLFTWG-EQN----- 211
QY 214 PPPSPGPGPTABAAALFSRECRPSAQLGALRVMLVWTT-AYFFLPFLCLSLYGLIGR 272
Db 212 -----RSADGQHQAGGLVCTPT-IHTATVKVVIQVNTFMFSFIPVWVISVLNTIIAN 261
QY 273 ELWSSRPL--RCPAAS-----GREGRHQTVRLVLLVVLAFIICWLPFHV 316
Db 262 KLTVMVRQAABEQGVCTVGGHEHTFSWAIBPGVQALRHGVRVLRVAVIAFVVCWLPYHV 321
QY 317 GRIIYINTEDSR---MMY-PSQYFNIVALQLFYLASINPILNLYSKYR-----AA 365
Db 322 RLMLFCYISDEQWTFDYDYHYFYVMVTVNLFVSVSTINPILNLYSANFRHIFLATLAC 381
QY 366 AFKLLARKSRPRGFHRSRT 386
Db 382 LCPVRRRRKRP-AFSRKADS 401

RESULT 3
JC7913
capa receptor (CG14575) - fruit fly (Drosophila sp.)
C;Species: Drosophila sp.
C;Date: 31-Mar-2003 #sequence_revision 31-Mar-2003 #text_change 31-Mar-2003
C;Accession: JC7913
R;Iversen, A.; Cazzamali, G.; Williamson, M.; Hauser, F., and Grimmelikhuijzen, C.J.P.
Biochem. Biophys. Res. Commun. 299, 628-633, 2002
A;Title: Molecular cloning and functional expression of a Drosophila receptor for the ne
A;Reference number: JC7913; MUID:22347021; PMID:12459185
A;Accession: JC7913
A;Molecule type: mRNA
A;Residues: 1-477 <IVE>
A;Cross-references: GB:AF505865
C;Comment: This receptor is a G-protein-coupled receptor stimulates renal (Maipighi)
```

```
C;Genetics:
A;Introns: 70/3; 121/2; 126/1; 218/2; 276/2; 301/2; 349/1; 402/3; 436/3

Query Match 21.0%; Score 453; DB 2; Length 477;
Best Local Similarity 29.7%; Pred. No. 1.8e-28;
Matches 117; Conservative 68; Mismatches 135; Indels 74; Gaps 11;

QY 8 SDGEGAREPPWPALPPCDERRCSP-----FPLGALVPVTVAVCLCLFVVGVS 55
Db 30 SDPSHGFEEDY-----ACGTFNCSPKEFVAVLGPQTLPLYKAVLITIFGGIFITGVG 85
QY 56 NVVTVMILIGRYDRMRTTNNLYLGSMAVSDLLILL-GLPPDLYLWRSRPPWVFPGLLCRLS 114
Db 86 NLLVCIVIRHSAMHTATNYVFLSAVSDLLYLLFGLTEVFLYMHQYDPDFGMPFKIR 145
QY 115 LYVGECTVATLLHMTALSVERVLAICRPLRARVLTTRVRVRLAIVLWALLSAGPFL 174
Db 146 AFISEACTYVSVEFTIVAFSMERFLAICHPHLHLYAMVGFKRAIRIITALTMTVSVISAI 204
QY 175 FLVGVEDQPGISVVPGLNGTARIASSPLSSPPLWLSRAPPPSPGPTAAEAAALFSRE 234
Db 205 -----GLLSDIQVLYNPLDHS-----RIEESAF----- 227
QY 235 CRPSAQLGALRVMLWVTTAYFPFLCLSLYGLIGRELWSSRRPLRG--PAASGRERG 292
Db 228 CSMSPKIVNEIPVFEVSFCIFFVPMILITLLYGRMGAKIRSRTNOKLGVQOQTNNRETR 287
QY 293 HRQ-----TVRLVVVVLAFIICWLPFHVGRITIIYNTEDSRMMYFSQYFNI-----VAL 341
Db 288 NSQMRKXTVIRMLAAVVIITFCWFFHLQRLIFYAKN-----MDNYLDINEALFSIAG 342
QY 342 QLFLYSASINPILNLYSKYRAAFAKLLLARKS 375
Db 343 FAYVSVCTVNPVSVMSRRYR-VAFRELLCGKA 375

RESULT 4
A88013
protein K10B4.4 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C;Accession: A88013
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C_eleg
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: A88013
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-418 <STO>
A;Cross-references: GB:chr_II; PIDN:AB71009.1; PID:g2429475; GSPDB:GN00020; CESP:K10B4.4
A;Note: similar to family I of G-protein coupled receptors
C;Genetics:
A;Gene: K10B4.4
A;Map position: 2

Query Match 20.3%; Score 437; DB 2; Length 418;
Best Local Similarity 30.1%; Pred. No. 2.9e-27;
Matches 109; Conservative 78; Mismatches 137; Indels 38; Gaps 10;

QY 29 RCSPPFLGALVPVTVAVCLCLFVVGSGNVVTVMLIGRYDRMRTTNNLYLGSMAVSDLL-I 87
Db 24 RCQ--SAGVIVPTVIITGYIFLLGLFGNICTCIVIAANKSMHNPNTNYLFLSAVSDIAL 81
QY 88 LLGLPPDLYR-LWRSRPPWFGPLLCRLSLYVGGCTVATLLHMTALSVERYLAICRPLRA 146
Db 82 ILGLPMFYSQSLDYSYPRFSEGCICARAFLEFTSYASIMICCFSEFWLAIChPLRS 141
QY 147 RVLTTRRRVRLAIVLWALLSAGPFLFVVGVEQDPP---GISVVPGLNGTARIASSPLA 203
Db 142 KIPSTLWRANVLIIIAWTISFVCALPIAFIVQINKLPLPEDAKYQPWTKVSPFAVGVLN 201
```


A;Reference number: I57949; MUID:94088493; PMID:8264565
A;Accession: I57949
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 341-363 <OC42>
A;Cross-references: GB:S67370; NID:G455947; PIDN:AA29371.1; PID:G455948
A;Experimental source: pituitary
R;Penetta, R.; Greenwood, M.; Patel, Y.C.
submitted to the EMBL Data Library, August 1993
A;Description: Correction of the nucleotide and amino acid sequence of the rat somatostatin receptor.
A;Reference number: S3244
A;Accession: S3244
A;Molecule type: mRNA
A;Residues: 309-363 <PEN>
A;Cross-references: EMBL:X74828; NID:G433911; PIDN:CAA52825.1; PID:G433912
C;Genetics:
C;Superfamily: vertebrate rhodopsin
Query Match 16.7%; Score 359; DB 2; Length 363;
Best Local Similarity 27.3%; Pred. No. 4.1e-21;
Matches 110; Conservative 60; Mismatches 149; Indels 84; Gaps 12;
QY 5 WNGSDGEGAREPPWALPCDERRCSPFLGA---LVPVAVCLCLFVVGSGNVVVM 61
DB 12 WNASAGSGNHN--WSLVG-----SASPMGARAVLPVLYLLVC--TVGLSGNTLVY 60
QY 62 LIGRYDRMRTTNLYLGSMVSDLLILLGLPFDLYLRWRSPWVFGPLLCRLSLYVGE 121
DB 61 VTLRHAKMTVTNVIILNADVLFMLGLFELATONAVVSYWPFSGFLCRLVNTLDGIN 120
QY 122 TYATLHMTALSVERYLAI CRPLARVLVTRRRVRLIALVMAVALLSAGPFLFLVGVEQ 181
DB 121 QFTSIFCLWMSVDVRLAVVHPTRSRARPRVAKMASAAVWFSLLMSLPLVFDVQE 180
QY 182 DPGISVVGGLNGTARIASSPLASSPPLWLSRAPPPSPGPETAEEAALFRCRPSAQ 241
DB 181 G-----WGTCNLS-----W-----PEPVG 194
QY 242 LGALRVMLWVTAYFFLPFLCLSLYGLI-----GRELWSSRRPLRGPAASGRERCHR 294
DB 195 LWGAFTYTSVLGFFGFLVLCYLLIVVKVKAAGRVGSSR-----RRSE 244
QY 295 QTVRVLLAVVLAFTICWLPFHVGRIIYIN--TEDSRMMYFSQYFNIVALQLFYLSASINPI 353
DB 245 KVRWVVVVVLFVGCWLPFFIVNIVNLAFTLPEEPTSGAGLYFVVVLS--YANSCANPL 302
QY 354 LYNLSKKYRAAFAFKLLARKSRPRGPHRSRDTAGEVAGDTGG 396
DB 303 LYGLSDNFRQSPRKVLCRLR---RGYGMEDADAIEPRDPKSG 341
RESULT 8
A46226
somatostatin receptor 3 - human
C;Species: Homo sapiens (man)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 24-Nov-1999
C;Accession: A46226; S32501
R;Yamada, Y.; Reisine, T.; Law, S.F.; Ihara, Y.; Kubota, A.; Kagimoto, S.; Seino, M.; Seino, M.; Endocrinol. 6, 2136-2142, 1992
A;Title: Somatostatin receptors, an expanding gene family: cloning and functional characterization.
A;Reference number: A46226; MUID:93149123; PMID:1337145
A;Accession: A46226
A;Molecule type: DNA
A;Residues: 1-418 <YAM>
A;Cross-references: GB:M96738; NID:G338498; PIDN:AAA60592.1; PID:G338499
A;Note: sequence extracted from NCBI backbone (NCBIN:123685, NCBIPI:123690)
R;Corness, J.D.; Demchyshyn, L.L.; Seeman, P.; van Tol, H.H.M.; Srikant, C.B.; Kent, G.; FEBS Lett. 321, 279-284, 1993
A;Title: A human somatostatin receptor. (SSTR3), located on chromosome 22, displays preference for [D-Phe1]somatostatin.
A;Reference number: S32501; MUID:93238970; PMID:8097479
A;Accession: S32501
A;Status: preliminary

A;Molecule type: DNA
A;Residues: 1-418 <COR>
C;Genetics:
A;Gene: GDB:SSTR3
A;Cross-references: GDB:134187; OMIM:182453
A;Map position: 22q13.1-22q13.1
A;Introns: #status absent
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane protein
F;44-70/Domain: transmembrane #status predicted <TM1>
F;81-106/Domain: transmembrane #status predicted <TM2>
F;118-139/Domain: transmembrane #status predicted <TM3>
F;159-181/Domain: transmembrane #status predicted <TM4>
F;203-233/Domain: transmembrane #status predicted <TM5>
F;255-282/Domain: transmembrane #status predicted <TM6>
F;289-316/Domain: transmembrane #status predicted <TM7>
F;17,30/Binding site: carboxylate (Asn) (covalent) #status predicted
F;116-191/Disulfide bonds: #status predicted
F;151,251,317,332/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted
F;251/Binding site: phosphate (Ser) (covalent) (by CAMP-dependent kinase) #status predicted
F;256/Binding site: phosphate (Thr) (covalent) (by CAMP-dependent kinase) #status predicted
F;412/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
Query Match 16.6%; Score 357.5; DB 2; Length 418;
Best Local Similarity 26.9%; Pred. No. 6.2e-21;
Matches 101; Conservative 63; Mismatches 146; Indels 65; Gaps 10;
QY 11 PEGAREPPWALPCDERRCSPFL-----GALVPTAVCLCLFVVGSGNVVTVMLIGR 65
DB 15 PENA-SSAWPPDPATLGNVAGSPAGLAVSGVLIPLVLYVC--VVGLLGNLSLVIYVVL 71
QY 66 YRDMRTTNLYLGSMVSDLLILLGLPFDLYLRWRSPWVFGPLLCRLSLYVGE 125
DB 72 HTASPSVTNVIILNADVLFMLGLPFLAAQNALSYWPFSGLMCLWAVDGINQPTS 130
QY 126 LHMATLSVERYLAI CRPLARVLVTRRRVRLIALVMAVALLSAGPFLFLVGVE 185
DB 131 IFCLTWMSVDVRLAVVHPTRSRARMTAPVARTVSAVAVVAVVLPVVVFSVG 184
QY 186 SVVPGGLNGTARIASSPLASSPPLWLSRAPPPSPGPETAEEAALFRCR 242
DB 185 -----PRGMST-----CHMQWPEPAAA 201
QY 243 GALRVMLWVTAYFFLPFLCLSLYGLIGRELWSSRRPLRGPAASGRERHROT 302
DB 202 WRAGFIYTAALGFFGFLVLCYLLIVVKVRSAGRRVWAPSCQRRRRSERRVTRMVA 261
QY 303 VLAFIICWLPFHVGRIIYINTE--DSRMMYFSQYFNIVALQLFYLSASINPI 361
DB 262 VVALFVLCWMPFYVNLVNVVCPLEPAPFFGLYFLVALP--YANSCANPLV 319
QY 362 YRAAFAFKLLARKSR 376
DB 320 FK-QGFRVLLRPSR 333
RESULT 9
I57955
somatostatin receptor - human
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Nov-1999
C;Accession: I57955
R;Panetta, R.; Greenwood, M.T.; Warszynska, A.; Demchyshyn, L.L.; Day, R.; Niznik, H.B.; Mol. Pharmacol. 45, 417-427, 1994
A;Title: Molecular cloning, functional characterization, and chromosomal localization of the human somatostatin receptor.
A;Reference number: I57955; MUID:94195267; PMID:7908405
A;Accession: I57955
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-363 <RES>
A;Cross-references: GB:I14865; NID:G431094; PIDN:AAA20828.1; PID:G431095
C;Genetics:
A;Gene: GDB:SST

A;Cross-references: GDB:119604; OMIM:182450
A;Map position: 3q28-3q28
C;Superfamily: vertebrate rhodopsin

Query Match 16.0%; Score 344; DB 2; Length 363;
Best Local Similarity 27.7%; Pred. No. 6.4e-20;
Matches 106; Conservative 58; Mismatches 131; Indels 88; Gaps 14;
QY 14 AREPPWPALPP-----CDERR-CSPPP-LGA-----LVPVTAVCLCLFVVGSGNVVTVML 62
DB 7 ASTPSWNASPGASGGDNRTLVGPAPAGARAVLPVLYLLVC--AAGLGGNTLVIYV 64
QY 63 IGRYDMRTTNLYLGSMAVSDLLILGLPFDLYRLWRSRPWPGPCLRLSLYVSGECT 122
DB 65 VLRFKMKTKTNIYILNLAVADVLYMLGLPP-LATQNAASFMPFGPVLCLRLVMTLDGVNQ 123
QY 123 YATLLHMTALSVRYLAICPLRLARVLVTRRRVRLALAVLWAVALLSAGPFLFLVGVQD 182
DB 124 FTSVFCUTVMSVDYLAIVHPLSSARWRPRVAKLASAAWVLSCLMSLPLLVFADVOE- 182
QY 183 PGISVVPGLNGTARIASSPLASSPPLMLSRAPPPSPGPETABAAALFSRECRPSPAOL 242
DB 183 -----GGTCNAS-----W-----PEPVL 196
QY 243 GALTVMVLTAVYFFLPFLCLSLYGLIGRELWSSRRPLRGPAAAG-----BERGHRQT 296
DB 197 WGAFFIYITAVLGFAPFLVLCYLLIVKV-----RAAGVRVGCVRSSRKV 246
QY 297 VRLLVVLVLAFLICWLPFHVGRIIYI-----NTEDSRMYFSQYFNIVALQLFVLSASIN 351
DB 247 TRMLVVLVLFAGCWLFPFTTNVNLAVLPQSPASAGLYP---FVVI---LSYANSCAN 300
QY 352 PILYNLISKYRAAFAKLLARK 374
DB 301 PVLGYFLSDNFRSQFVKVCLRK 323

RESULT 10
JN0763
somatostatin receptor 5 - human
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 20-Jun-2000
C;Accession: JN0763
R;Yamada, Y.; Kagimoto, S.; Kubota, A.; Yasuda, K.; Masuda, K.; Someya, Y.; Ihara, Y.; I
Biochem. Biophys. Res. Commun. 195, 844-852, 1993
A;Title: Cloning, functional expression and pharmacological characterization of a fourth
A;Reference number: JN0762; MUID:93384611; PMID:8373420
A;Accession: JN0763
A;Molecule type: DNA
A;Cross-references: DDBJ:D16827; NID:9487683; PIDN:BAA04107.1; PID:9487684
C;Comment: This protein is a member of somatostatin receptor family.
C;Genetics:
A;Gene: GDB:SSTR5
A;Cross-references: GDB:138452; OMIM:182455
A;Map position: 16p13.3-16p13.3
A;Introns: #status absent
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; glycoprotein; lipoprotein; phosphoprotein; thiol
F;40-66/Domain: transmembrane #status predicted <TM1>
F;77-102/Domain: transmembrane #status predicted <TM2>
F;114-135/Domain: transmembrane #status predicted <TM3>
F;155-177/Domain: transmembrane #status predicted <TM4>
F;196-228/Domain: transmembrane #status predicted <TM5>
F;246-273/Domain: transmembrane #status predicted <TM6>
F;280-307/Domain: transmembrane #status predicted <TM7>
F;13,26,187/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;112-186/Diulfide bonds: #status predicted
F;242,325/Binding site: phosphate (Ser) (covalent) (by CAMP-dependent kinase) #status pr
F;247/Binding site: phosphate (Thr) (covalent) (by CAMP-dependent kinase) #status predic
F;320/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match 16.0%; Score 344; DB 2; Length 364;

Best Local Similarity 27.7%; Pred. No. 6.4e-20;
Matches 106; Conservative 58; Mismatches 131; Indels 88; Gaps 14;
QY 14 AREPPWPALPP-----CDERR-CSPPP-LGA-----LVPVTAVCLCLFVVGSGNVVTVML 62
DB 7 ASTPSWNASPGASGGDNRTLVGPAPAGARAVLPVLYLLVC--AAGLGGNTLVIYV 64
QY 63 IGRYDMRTTNLYLGSMAVSDLLILGLPFDLYRLWRSRPWPGPCLRLSLYVSGECT 122
DB 65 VLRFKMKTKTNIYILNLAVADVLYMLGLPP-LATQNAASFMPFGPVLCLRLVMTLDGVNQ 123
QY 123 YATLLHMTALSVRYLAICPLRLARVLVTRRRVRLALAVLWAVALLSAGPFLFLVGVQD 182
DB 124 FTSVFCUTVMSVDYLAIVHPLSSARWRPRVAKLASAAWVLSCLMSLPLLVFADVOE- 182
QY 183 PGISVVPGLNGTARIASSPLASSPPLMLSRAPPPSPGPETABAAALFSRECRPSPAOL 242
DB 183 -----GGTCNAS-----W-----PEPVL 196
QY 243 GALTVMVLTAVYFFLPFLCLSLYGLIGRELWSSRRPLRGPAAAG-----BERGHRQT 296
DB 197 WGAFFIYITAVLGFAPFLVLCYLLIVKV-----RAAGVRVGCVRSSRKV 246
QY 297 VRLLVVLVLAFLICWLPFHVGRIIYI-----NTEDSRMYFSQYFNIVALQLFVLSASIN 351
DB 247 TRMLVVLVLFAGCWLFPFTTNVNLAVLPQSPASAGLYP---FVVI---LSYANSCAN 300
QY 352 PILYNLISKYRAAFAKLLARK 374
DB 301 PVLGYFLSDNFRSQFVKVCLRK 323

RESULT 11
A55259
kappa opioid receptor - guinea pig
N;Alternate names: dynorphin receptor
C;Species: Cavia porcellus (guinea pig)
C;Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 24-Nov-1999
C;Accession: A55259
R;Xie, G.; Meng, F.; Mansour, A.; Thompson, R.C.; Hoversten, M.T.; Goldstein, A.; Watson,
Proc. Natl. Acad. Sci. U.S.A. 91, 3779-3783, 1994
A;Title: Primary structure and functional expression of a guinea pig kappa opioid (dynor
A;Reference number: A55259; MUID:94224825; PMID:8170987
A;Accession: A55259
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-380 <XIE>
A;Cross-references: GB:U04092; NID:9476106; PIDN:AAA67171.1; PID:9476107
C;Superfamily: vertebrate rhodopsin
C;Keywords: transmembrane protein
Query Match 15.9%; Score 343.5; DB 2; Length 380;
Best Local Similarity 25.1%; Pred. No. 7.3e-20;
Matches 102; Conservative 74; Mismatches 147; Indels 83; Gaps 12;
QY 2 GSPW-----NGSDGPGAREPP---WPALPPCDERRCSPPLGALVPVTVAVCLCLF 49
DB 26 GSAWLPQWAEPCDNGSAGPQDEQLEPAHISPAIP-----VIIITAVYSVP 70
QY 50 VVGSGNVTVTVMLIGRYDMRTTNLYLGSMAVSDLLILGLPFDLYRLWRSRPWPGPCL 109
DB 71 VUGLVGNSLVMPFVIIRYTKMTATNIYIFNLADALAVTTTTPFQ-STVTLMSNSWPGDV 129
QY 110 LCRSLYVVGSGCTYATLLHMTALSVRYLAICPLRLARVLVTRRRVRLALAVLWAVALLS 169
DB 130 LCKIVISIDYNNMFTSIFTLTMMVSDRYIAVCHPVKALDPRFLKAKIINICWLLSSV 189
QY 170 AGPFLVGVQDQPGISVVPGLNGTARIASSPLASSPPLMLSRAPPPSPGPETABAAA 229
DB 190 GISAILGGTKVREDVDIIE-----CSLQPDDDYSWD----- 223
QY 230 LFSRECRPSPAQLGALTVMVLTAVYFFLPFLCLSLYGLIGRELWSSRRPLRGPAAAG 289

Db 224 LFWKIC-----VFVFAFVPIVLLIIVCYTLMILRL-KSVALLSG--SREK 265

QY 290 ERGRHQTIVRLVLLVLAFLIICWLPFHVGRIIYI--NTEDSRMMVFSQYFNIVALQLFYLS 347

Db 266 DNLRLRITRLVLLVAVFICWPIFIHIFILVEALGSTSHSTAALSSYF---CIALGYTN 322

QY 348 ASINPILYNLISKYRAA-----AFKLLARKSRPGRPHRSRDTA 387

Db 323 SSLNPILYALDENFKRCDFCFPIKMRMERQSTSRVNTVQDPA 368

RESULT 12

JE0296

thyrotropin releasing hormone receptor - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 21-Jul-2000

C;Accession: JE0296

R;Iwadani, H.; Nakamura, T.; Itoh, J.; Iwase, H.; Kanatani, A.; Borkowski, J.; Ihara, M. Biochem. Biophys. Res. Commun. 250, 68-71, 1998

A;Title: Cloning and characterization of a new subtype of thyrotropin-releasing hormone A;Reference number: JE0296; MUID:98407892; PMID:9735333

A;Accession: JE0296

A;Molecule type: mRNA

A;Residues: 1-352 <ITA>

A;Cross-references: DDBJ:AB015645; NID:G3660553; PIDN:BAA33437.1; PID:G3660554

C;Superfamily: adenosine receptor A1

F;26-48/Domain: transmembrane #status predicted <TM1>

F;58-80/Domain: transmembrane #status predicted <TM2>

F;97-118/Domain: transmembrane #status predicted <TM3>

F;142-165/Domain: transmembrane #status predicted <TM4>

F;188-209/Domain: transmembrane #status predicted <TM5>

F;252-273/Domain: transmembrane #status predicted <TM6>

F;282-304/Domain: transmembrane #status predicted <TM7>

Query Match 15.7%; Score 339; DB 2; Length 352;

Best Local Similarity 28.8%; Pred. No. 1.5e-19;

Matches 108; Conservative 52; Mismatches 109; Indels 106; Gaps 13;

QY 41 VTAVCLCLFY--VGSGNVVTVMLIGRYDMRTTNNLYLGSMAVSDLLILL--GLPPDLY 96

Db 23 VVSFLVLLVCTIGVGNAMVILVLTSDMHTPTNCYLVSLADLLVLAAGLPNVSD 82

QY 97 RLWRSRPWFGLCLRLSLXVGGCTYATLL-----HMTLSVERYLAICRPLRARVL 149

Db 83 SL--VGHWYGRAGCL-----GITVQYGLINVSFSILFTVERYIAICPLRAQTV 133

QY 150 VTRRRVRLTAVLWVALLSAGFLFLVG-----EQDPSVVPGLNGTARIASPLA 203

Db 134 CTVARAKRIIAGVGTSLYCLLWFLVLDLNRDNRLEGYKVPRL-----181

QY 204 SSPPLMLSRAPPSPSPGPETAEEALFSRECRPSAQLGALRVMLVWTTTAYFFLPCLCL 263

Db 182 -----YLPYLLDFAVFIIGPLVLT 201

QY 264 SILYGLIGR-----ELWSSRRPLRG--PAASGR-----ERGHQTIVRLVLLVLAFLI 308

Db 202 LVLYGLIGRILFQSPLSQEAQWQERQPHGQSEAAAGNCSRAKSRKQATEMLAVVLLFA 261

QY 309 ICWLPFHVGRIIYI--NTEDSRMMVFSQYFNIVALQLF-----YLSASINPILYNLISKYR 363

Db 262 VLATPYRT--LVLLNS-----FVAQPLDPWLLFCRTCVTNSAVNPVYSLMSQKR 313

QY 364 AAAPKLLARKSRPR 378

Db 314 AAFKLUCWCRAGFPQ 328

RESULT 13

S36143

kappa opioid receptor - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 10-Dec-1993 #sequence_revision 19-Oct-1995 #text_change 20-Jun-2000

C;Accession: S36143; S38825; S36102; S39015; A48789

R;Nishi, M.; Takeshima, H.; Fukuda, K.; Kato, S.; Mori, K. FEBS Lett. 330, 77-80, 1993

A;Title: cDNA cloning and pharmacological characterization of an opioid receptor with high affinity for morphine; A;Reference number: S36143; MUID:93380575; PMID:8396539

A;Accession: S36143

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-380 <NIS>

A;Cross-references: GB:D16534; NID:G409390; PIDN:BAA03971.1; PID:G415310

R;Chen, Y.; Westek, A.; Liu, J.; Yu, L. Biochem. J. 295, 625-628, 1993

A;Title: Molecular cloning of a rat kappa opioid receptor reveals sequence similarities to the mu and delta receptors; A;Reference number: S38825; MUID:94059008; PMID:8240267

A;Accession: S38825

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-380 <CHE>

A;Cross-references: GB:I22001; NID:G409236; PIDN:AAA1495.1; PID:G409237

R;Minami, M.; Toya, T.; Katao, Y.; Maekawa, K.; Nakamura, S.; Onogi, T.; Kaneko, S.; Satoh, T. FEBS Lett. 329, 291-295, 1993

A;Title: Cloning and expression of a cDNA for the rat kappa-opioid receptor. A;Reference number: S36102; MUID:93374033; PMID:8103466

A;Accession: S36102

A;Molecule type: mRNA

A;Residues: 1-41, 'L', 43-380 <MIN>

A;Cross-references: GB:D16829; NID:G404115; PIDN:BAA04109.1; PID:G404116

R;Li, S.; Zhu, J.; Chen, C.; Chen, Y.W.; Deriel, J.K.; Ashby, B.; Liu-Chen, L.Y. Biochem. J. 295, 629-633, 1993

A;Title: Molecular cloning and expression of a rat kappa opioid receptor. A;Reference number: S39015; MUID:94059009; PMID:8240268

A;Accession: S39015

A;Molecule type: mRNA

A;Residues: 1-344, 'Y', 346-380 <LIS>

A;Cross-references: GB:I22536; NID:G425188; PIDN:AAA1496.1; PID:G425189

R;Meng, F.; Xie, G. Proc. Natl. Acad. Sci. U.S.A. 90, 9954-9958, 1993

A;Title: Cloning and pharmacological characterization of a rat kappa opioid receptor. A;Reference number: A48789; MUID:94052210; PMID:8234341

A;Accession: A48789

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-380 <RES>

A;Cross-references: EMBL:U00442; NID:G403486; PIDN:AAA18261.1; PID:G403487

C;Superfamily: vertebrate rhodopsin

C;Keywords: G protein-coupled receptor; transmembrane protein

Query Match 15.7%; Score 339; DB 2; Length 380;

Best Local Similarity 25.5%; Pred. No. 1.7e-19;

Matches 103; Conservative 75; Mismatches 146; Indels 80; Gaps 13;

QY 6 NGSDDGPEGAREPP--WPALPPCDERRCSPFPLGALVPVTAIVCLCLFVVGSGNVTVML 62

Db 39 NGSVSGEDQOLEPAHISPAIP-----VIITAVSVVVFVGLVGNLSVWFV 83

QY 63 IGRYDMRTTNNLYLGSMAVSDLLILLGLPFDLYRLRSRPWFGLCLRLSLXVGGCT 122

Db 84 IIRYTKTKATNIYIENLADALVTTTHFPQ--SAVYLMNSWPFGLVCLKIVISIDYNN 142

QY 123 YATLHMTALSVRYLAICRPLRARVLVTRRRVRLIIVLWVALLSAGFLFLVGVEQD 182

Db 143 FTSIFTLTMSVDRYTAVCHPKALDFRTPAKIINICIWLLASSVGISAIVLGGTKVR 202

QY 183 PGISVVPGLNGTARIASSPLASSPPLWLSRAPPSPPGPETAEEALFSRECRPSAQL 242

Db 203 EDVDVIE-----CSLOFPDDEYSWMD-----LFWKIC-----229

QY 243 GALTIVMLVTTTAYFFLPCLCLSLYGLIGRELWSSRRPLRGPAASGRGHQTIVRLV 302

Db 230 -----VFVFAFVPIVLLIIVCYTLMILRL-KSVALLSG--SREKDLNRITKLVLV 278

QY 303 VLAFIICWLPFHVGRIIYI--NTEDSRMMVFSQYFNIVALQLFYLSASINPILYNLISK 360

Db 279 VVAVFIICWPIFIHIFILVEALGSTSHSTAALSSYF---CIALGYTNSSLNPVLYAFLDE 335

Qy 123 YATLLHWTALSVERYLAICRPLRARVLVTRRRVRALLIAVLWAVALLSAGPFLFVLGVSD 189
Dd :
Db 143 FTSITLTTMMSVDRIYAVCPVKALDPRTPLKAKIINICIWLLSSSGVISAILVGCTKVR 202
Qy 183 PGISVVPGINGTARIASSPLASSPLMLSRAPPPSPGSPETAEAAALFSRECRPSAQL 242
Dd :
Db 203 EDVDVIE-----CSLQPDDDDYSWD-----LFMKIC----- 229
Qy 243 GALLRVLMWTVTYFFLPFLCLSLILGYIGRELWSRRRLRGPAASGRGRHRQTVRVLV 302
Dd :
Db 230 -----VPFAPVIPVLIIVCYTLMILRL-KSVRLLSG--SREKDRNLRRITRLVLV 278
Qy 303 VVAFTICWLPHFHVGRIIYI--NTEDSRMMVFYSFYFNIVALQLPYLSASINPILYNLSK 360
Dd :
Db 279 VWAFVVCMTPIHFILVEALGSTSHSTAALSSYYF---CIALGVTNSSLNPILYAFLDE 335
Qy 361 KYRAA-----AFKLLARKSRPRGFHSRDTA 387
Dd :
Db 336 NFKRCFRDPCFPCLKRMEROSTSRVNTVQDFA 368

RESULT 15
JN0605
somatostatin receptor 4 - human
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence revision 30-Sep-1993 #text_change 20-Jun-2000
C:Accession: JN0605; JN0762; A47457
R:Xu, Y.; Song, J.; Bruno, J.F.; Berelowitz, M.
Biochem. Biophys. Res. Commun. 193, 648-652, 1993
A:Title: Molecular cloning and sequencing of a human somatostatin receptor, hSR
A:Reference number: JN0605; MUID:93290656; PMID:8512564
A:Accession: JN0605
A:Molecule type: DNA
A:Residues: 1-388 <Y>
A:Cross-references: GB:L14856; NID:G9292499; PIDN:AAA36623.1; PID:G9292500
R:Yamada, Y.; Kagimoto, S.; Kubota, A.; Yasuda, K.; Masuda, K.; Someya, Y.; Iba
Biochem. Biophys. Res. Commun. 195, 844-852, 1993
A:Title: Cloning, functional expression and pharmacological characterization o
A:Reference number: JN0762; MUID:93384611; PMID:8373420
A:Accession: JN0762
A:Molecule type: DNA
A:Residues: 1-388 <YAM>
A:Cross-references: GB:D16826; NID:G693907; PIDN:BAA04106.1; PID:G693908
R:Roher, L.; Raulf, F.; Bruns, C.; Buettner, R.; Hofstaedter, P.; Schule, R.
Proc. Natl. Acad. Sci. U.S.A. 90, 4196-4200, 1993
A:Title: Cloning and characterization of a fourth human somatostatin receptor.
A:Reference number: A47457; MUID:93248256; PMID:8483934
A:Accession: A47457
A:Molecule type: DNA
A:Residues: 1-82, 'T', 84-364, 'K', 366-388 <ROH>
A:Cross-references: GB:L07833; NID:G107429; PIDN:AAA60565.1; PID:G307430
A>Note: sequence extracted from NCBI backbone (NCBIN:130856, NCBI:P:130858)
C:Comment: This protein mediates the diverse actions of the tetradecaptide som
C:Genetics:
A:Gene: GDB:SSTR4
A:Cross-references: GDB:202662; OMIM:182454
A:Map position: 20p11.2-20p11.2
A:Introns: #status absent
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; lipopr
P:F47-73/Domain: transmembrane #status predicted <TM1>
P:F84-109/Domain: transmembrane #status predicted <TM2>
P:F121-142/Domain: transmembrane #status predicted <TM3>
P:F162-184/Domain: transmembrane #status predicted <TM4>
P:F208-238/Domain: transmembrane #status predicted <TM5>
P:F257-284/Domain: transmembrane #status predicted <TM6>
P:F291-314/Domain: transmembrane #status predicted <TM7>
P:F24/Binding site: carbohydrate (asn) (covalent) #status predicted
P:F119-198/Disulfide bonds: #status predicted
P:F161,253/Binding site: phosphate (Ser)(covalent) (by CAMP-dependent kinase) f
P:F327/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match	15.5%;	Score 335;	DB 2;	Length 388;	
Best Local Similarity	26.9%;	Pred. No. 3.6e-19;			
Matches 105;	Conservative	65;	Mismatches 128;	Indels 92; Gaps 15;	
QY	1	MGSPW----	NGSDGPEGAREPPWPA	PPCDERRCSPPLGALVPVTAVCLCLFVWGSVN 56	
DB	15	LGTWPSAANA	SAPAEAE-----	AVAGFGDAR-----AAGWAIQCIYALVCLVGLVGN 65	
QY	57	VVTVMILIGRYDMRTT	NLYGSMVSDLLILGL	PPDLYRLWRSRPPWFGPCLLCRLUSLY 116	
DB	66	ALVIFVILRYAKMKTAT	NIYLLLAVALDELFL	MSVPP-VASSAALRHWPFSGVLCRAVL 124	
QY	117	VBEGCCTYATLLHMTAL	SVERYLAI	CRPLRARVLVTRER--VRALIY-LNAVALLSAGPF 173	
DB	125	VDGLNMFTSVFC	UTVLVSVDYRVA	VUHPLEA--ATYRPSVAKLILNGVWLASLVLTLPI 181	
QY	174	LFLVGVGEQD	FGISVWFGNLGTARI	ASSPLMLSRAPPPSPSGPETA	AAALFSR 233
DB	182	AIFADTR-----	FARGGQA--	VACNLQWHPAM-----	SAVF-- 211
QY	234	ECRPSPAQLGAL	RYMLWVTYATFEL	PLCLSLYGLIGREL-----	WSSRRPLRGPA 285
DB	212	-----	VVYTFLLGFLP	VIATGLCYLLIVG	KORAVALRAGWQRR----- 251
QY	286	ASGRERGHRTV	RLVLLVWVLAFTIC	MLPPHVGRH--IYINTE	DSRMMYFQGYENIVALQL 343
DB	252	-----	RSEKTI	FLVLMVVVVFVLC	WMPFFVYQVLLNLVVTSLDATV-----NHVSLIL 299
QY	344	FYLSASINPILYN	LISKRYAAAFKLLAR		373
DB	300	STANSCANPILY	GLFSLDNFRSRQ	RVLCRLR 329	

Search completed: January 1, 2004, 06:47:44
Job time : 28.3308 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 1, 2004, 06:19:59 ; Search time 67.6341 Seconds
(without alignments)
1571.954 Million cell updates/sec

Title: US-09-719-485-3

Perfect score: 2155

Sequence: 1 MGSPWNGSDGPEGAREPPWP.....DTGGDTVGTETTSANVKTWG 412

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 23.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phage.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_rvirus.*

16: sp_bacteriap.*

17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	907	42.1	363	13	Q93413
2	861.5	40.0	364	11	Q8BWX8
3	851	39.5	374	13	Q93412
4	825.5	38.3	366	6	Q8MH25
5	608	28.2	559	13	Q93414
6	487.5	22.6	426	4	Q8NE20
7	485.5	22.5	426	4	Q9HB89
8	478.5	22.2	395	11	Q91276
9	478.5	22.2	395	11	Q8BZ39
10	477.5	22.2	395	11	Q9JIB1
11	477	22.1	415	4	Q96AM5
12	476	22.1	415	4	Q9GZ04
13	475.5	22.1	395	11	Q9ESQ4
14	457.5	21.2	405	11	O55040
15	453	21.0	477	5	Q8ITC7
16	437	20.3	418	5	O17239

Q9VFN4 drosophila
Q8ITC9 drosophila
Q8ITD0 drosophila
Q9VFW6 drosophila
Q9JIB2 rattus norv
Q9JIB5 rattus norv
Q8VIF5 mus musculus
Q920Q5 mus musculus
Q9VW5 drosophila
Q8ITD1 drosophila
Q18701 caenorhabdi
Q9JK40 mus musculus
Q8JFZ6 xenopus lae
Q9DDR0 xenopus lae
Q8JID5 carassius a
Q8MI04 ovis aries
Q8UW15 fugu rubrip
Q8GQ4 carassius a
Q97914 ovis aries
Q8820 rattus norv
Q9R297 rattus norv
Q9DDR1 xenopus lae
Q9DFA9 catostomus
Q8IWP3 homo sapien
Q8JFZ7 xenopus lae
Q9QW3 rattus norv
Q9ERT2 mus musculus
Q8JFZ5 xenopus lae
Q8IY58 homo sapien

ALIGNMENTS

RESULT 1

Q93413 ID O93413 PRELIMINARY; PRT; 363 AA.
AC O93413;
DT 01-NOV-1998 (T-EMBLrel. 08, Created)
DT 01-NOV-1998 (T-EMBLrel. 08, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Orphan G protein-coupled receptor.
OS Spherooides nepheus.
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Spherooides.
OX NCBI_TaxID=39110;
RN [1]
RP MEDLINE=20092336; PubMed=10628755;
RA Palya O.C., Feighner S.D., Tan C.P., McKee K.K., Hreniuk D.L.,
Gao Y.D., Schleim K.D., Yang L., Morriello G.J., Nargund R.,
Pattett A.A., Howard A.D., Smith R.G.;
RT "Ligand activation domain of human orphan growth hormone (GH)
secretagogue receptor (GHS-R) conserved from pufferfish to humans.";
RL Mol. Endocrinol. 14:160-169(2000).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AF082210; AAC33473.1; -.
DR InterPro; IPR000276; GPCR_Rhodopn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS00262; G-PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane.
SQ SEQUENCE 363 AA; 41324 MW; 74518BD9CD310991 CRC64;

Query Match 42.1%; Score 907; DB 13; Length 363;

Best Local Similarity 50.6%; Pred. No. 4.7e-67;

Matches 180; Conservative 56; Mismatches 76; Indels 44; Gaps 5;

OY 31 SPFFLGALVPVTAACLCLFVVGVSGNVVTVMLIGRYDRMTTTLNLYLGSNAVSDLLILG 90

```
Db 31 SLEPASTLIPVTICILIFVVGVTGNTWTIIIOYFKDMKTTNLYLSSMAVSDLVFLC 90
QY 91 LPDLYELWRSRPVFGPLCLRLSLVYEGCTVATLHMTALSVERVLAICRPLRARVLY 150
Db 91 LPDLYELWRSRPVFGPLCLRLSLVYEGCTVATLHMTALSVERVLAICRPLRARVLY 150
QY 151 TRRRVRLIALVMAVALLSAGPFLVLGVGEODPGISVVVGLNGTARIASSPLASSPPLWL 210
Db 151 TRRRVRLIALVMAVALLSAGPFLVLGVGEODPGISVVVGLNGTARIASSPLASSPPLWL 210
QY 211 SRAPPPSPGPGPETAFAALFSRECRPSPAQL--GALRVMLWVTYATFFLPFLCLSLYLG 268
Db 186 ----HPDYNTG-----QCKTGTVAISSGQLHIMWSTVTFPCPMLCLFLYG 229
QY 269 LIGRELWSSRPLRGPAAASGRCHROTVRVLLVVLAFIICWLPFHVGHRIIYNTEDSR 328
Db 230 SIGCKLWKSNDQGPALARSERSHRTVKILVVVVLAFIICWLPFHVGHRIIYNTEDSR 289
QY 329 MMVFSQYFNIVALQLFYLASINPILYNLSKRYAAAFKLLARKSRPRGFHRSR 384
Db 290 TAMSQNFNWSMVLVCLYSINPVVYNLSKRYAAAFKLLARKSRPRGFHRSR 344

RESULT 2
Q8BWX8 PRELIMINARY; PRT; 364 AA.
AC Q8BWX8
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Growth hormone secretagogue receptor type 1 homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR ENBL; AK049871; BAC33866.1; -.
SQ SEQUENCE 364 AA; 40969 MW; 8F1214E58EF3B2E8 CRC64;

Query Match 40.0%; Score 861.5; DB 11; Length 364;
Best Local Similarity 48.3%; Pred. No. 2,7e-63;
Matches 185; Conservative 52; Mismatches 81; Indels 65; Gaps 11;

QY 5 WNGSDGEGAREP-----PWPALPPCD---ERRCSFPPLGALVPVTVAVCLFLVVGVS 55
Db 2 WNAT--PSEPEPENVTLDLDDWDASPGNDSLSDELPLFPAPLAGVTATCVAFVVGISG 59
QY 56 NVVTVMILGVRDMRTTNLYLCSMAVSDLLILGLPFDLYRLWRSRPWFGPLCLRLSL 115
Db 60 NLUTLVVSRFRRETTNLYLCSMAFSDLLIFLCMPDLVRLWQVRPWNFGDLCKLFQ 119
QY 116 YVGECTYATLLHMTALSVERVLAICRPLRARVLTTRRRVRLIALVMAVALLSAGPFLF 175
Db 120 FVSECTYATVLTALSVERVLAICRPLRAKVVTGKRVKLVILVIAVAFCSAGPIFV 179
QY 176 LVGVEODPGISVVVGLNGTARIASSPLASSPPLWLRAPPPSPGETAFAALFSRECR 235
Db 180 LVGVEHE-----NGT-----DPRD-----TNEC 197
QY 236 RPS--PAQLGALRVMLWVTYATFFLPFLCLSLYLGILGRELWSSRPLRGPAAAG--RE 290
Db 198 RATEFAVRSGLLTMMWVSSVFFLPVFCVLTVLSIGRLKW--RR--RGDAAVGSLRD 253
QY 291 RGHROTVRVLLVVLAFIICWLPFHVGHRIIYNTEDS---RMMVFSQYFNIVALQLFYL 347
```

```
Db 254 QNHKQTVKMLVAVVVPFAFILCWLPFHVGRVLFSSKSPGSLIAQISOYCNLVSFVLFLY 313
QY 348 ASINPILYNLSKKYRAAFAKLL 370
Db 314 AAINPILYNLSKKYRAVAVFKLL 336

RESULT 3
Q93412 PRELIMINARY; PRT; 374 AA.
ID Q93412
AC Q93412;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Orphan G protein-coupled receptor.
OS Spherooides nephelus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Spherooides.
OX NCBI_TaxID=39110;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=20092336; PubMed=10628755;
RA Palyha O.C., Feighner S.D., Tan C.P., McKee K.K., Hreniuk D.L.,
RA Gao Y.D., Schleim K.D., Yang L., Morriello G.J., Nargund R.,
RA Patchett A.A., Howard A.D., Smith R.G.;
RT "Ligand activation domain of human orphan growth hormone (GH)
RT secretagogue receptor (GHS-R) conserved from pufferfish to humans.";
RL Mol. Endocrinol. 14:160-169(2000).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR ENBL; AF082209; AAC33472.1; -.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR000524; HTH_GntR.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PROSITE; PS00237; G PROTEIN RECEPTOR FL1; 1.
DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL2; 1.
DR PROSITE; PS00043; HTH_GNTR_FAMILY; 1.
DR KW G-protein coupled receptor; Receptor; Transmembrane.
SQ SEQUENCE 374 AA; 42324 MW; 2CF9304F004C7A16 CRC64;

Query Match 39.5%; Score 851; DB 13; Length 374;
Best Local Similarity 45.5%; Pred. No. 2.1e-62;
Matches 187; Conservative 65; Mismatches 101; Indels 58; Gaps 12;

QY 5 WNGS-DGEGAREPWPALPPCDERRCSFPPLGALVPVTVAVCLFLVVGVSNNVTVMIL 63
Db 13 WEGSHNGTAGLELPP-----LNYISIPLLAVITVACTVLTGVGVNMTILVV 61
QY 64 GRYDRMTTNLYLCSMAVSDLLILGLPFDLYRLWRSRPWFGPLCLRLSLYVGGCTY 123
Db 62 SRYDRMTTNLYLCSMAVSDLFIFVCMPLDLYRMWRYRPRFGDLCFLQFVSSECTY 121
QY 124 ATLLHMTALSVERVLAICRPLRARVLTTRRRVRLIALVMAVALLSAGPFLVGVGEODP 193
Db 122 STILCTALSVERVLAICRPLRAKALVTKRRAKALLTLLMTVLSLSAGPFWVNVGKDS 181
QY 184 GTSVVGPLNGTARIASSPL--ASSPPLWLRAPPPSPGETAFAALFSRECRPS--PA 240
Db 182 --IMFEN-----SSDLNNESSWPL-----EAVDTRECRMTQYAV 212
QY 241 QIGALRVMLWVTYATFFLPFLCLSLYLGILGRELWSSRPLR--GPAASGRGRHROTVR 299
Db 213 ESGLMEANWLVSSVFFPFCVLTVLYGLIGRLMLRHRETTINSRVAYRDKSNROTQM 272
QY 300 LNVVLAFLIICWLPFHVGHRIIYNTEDS---RMMVFSQYFNIVALQLFYLASINPILYN 356
Db 273 LVVVLAFLVCLWLPFHVGRVLFQRLSDAPSLLSLSSVSVVLFVLSAAINPILYN 332
QY 357 LISKYRAAAAFKLLARKSRPRGFHRSRDTAGEVAGDTGGDTVGYTETSAN 407
```


Db 333 TMSWKYRGAVARLFGVSDSPPO---RGR-TASTVKMD-----GWTESTVS 373

RESULT 4

Q8MHZ5 PRELIMINARY; PRT; 366 AA.
AC Q8MHZ5
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Ghrelin/growth hormone secretagogue receptor.
OS Ovis aries (Sheep).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1, 2, and 3; TISSUE=Pituitary;
RA Murata T., Ikegami R., Morita Y., Shinozaki K.;
RT "Sheep Ghrelin/Growth Hormone Secretagogue Receptor";
RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
DR ENBL; AY093948; AAM19733.1;
DR ENBL; AY093949; AAM19734.1;
DR ENBL; AY093950; AAM19735.1;
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR004704; IID_comp.
DR Pfam; PF00001; 7tm.1; 1.
DR Pfam; PF03613; EIID-AGN; 1.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 366 AA; 41486 MW; 2F276651BC6C5D57 CRC64;

Query Match 38.3%; Score 825.5; DB 6; Length 366;
Best Local Similarity 42.3%; Pred. No. 2.6e-60;
Matches 176; Conservative 62; Mismatches 113; Indels 65; Gaps 8;

QY 5 WNGSDGPE---GAREPPPALPPCD---ERRCSPEPLGALVPVTAVALCLFVVGSGN 56
DB 2 WNAIRSEELGNLTLPLDLDWAAPDNDLSLDELPLFPALLAGVATATCVALFVVGAGN 61
QY 57 VVTVMILGRYDMRTTNLVLGSMVSDLLILGLPDLVLRWRSRPWFGPLLCRLSL 116
DB 62 LTLVLRVSRRELRTTNLVLSSNAFSDLLIFLCMLDVLVLRWYRPNWGLDLCKLFQF 121
QY 117 VSGCTATLHMTALSVRYLAICRPLRVLVTRRRRRLAIVLWAVALLSAGPFLFL 176
DB 122 VSESCTYASVLITATLSVRYFAICFPLRAKVITKGRVLAIVLWAVAFCSAWPIFML 181
QY 177 VGYEQDPCISVPLGNTARIASSPLASSPLMSRAPPPSPGPTAAALFSECR 236
DB 182 VGYEHE-----NGT-----DPRD-----TNECR 199
QY 237 PS---PAQLGALRVMLVTTAYFFLPFLCLSLYLGLIGRELWSSRRPLRGAASGRGRHR 294
DB 200 ATPEAVRSGLTINWVSSIFFFLPVCLTVLYSLIGRLKWRRRSRVVGASLURDNHK 259
QY 295 QTVRLVLVVLAFICLWLPFHVGRIIYINTEDS---RMVFSQYFNIVALQLFYLASIN 351
DB 260 QTVKMLAVVFAFVLCWLPFHVGRLYFSKSPGVSVEIAQISQYCNLVSVFLFFYSAIN 319
QY 352 PILYNLSKKYRAAFKLLARKSRPRGRFHSRRTAGEVAGDTCGDVTGVTTESAN 407
DB 320 PILYNLSKKYRAVAVFKLLGPFPSQKSLTLKDESSR-----AWTESSIN 365

RESULT 5

Q93414 PRELIMINARY; PRT; 559 AA.
ID Q93414
AC Q93414;

DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Orphan G protein-coupled receptor.
OS Spherooides nephelus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Spherooides.
OX NCBI_TaxID=39110;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20092336; PubMed=10628755;
RA Palyha O.C., Feighner S.D., Tan C.P., McKee K.K., Hreniuk D.L.,
RA Gao Y.D., Schleich K.D., Yang L., Morriello G.J., Nargund R.,
RA Patchett A.A., Howard A.D., Smith R.G.;
RT "Ligand activation domain of human orphan growth hormone (GH)
secretagogue receptor (GHS-R) conserved from pufferfish to humans.";
RL Mol. Endocrinol. 14:160-169(2000).
DR ENBL; AF082211; AAC33474.1;
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 559 AA; 61956 MW; AFB9F35FAFOCD5F6 CRC64;

Query Match 28.2%; Score 608; DB 13; Length 559;
Best Local Similarity 29.7%; Pred. No. 4.3e-42;
Matches 156; Conservative 65; Mismatches 124; Indels 180; Gaps 12;

QY 22 LPPCDERRC---SPFPLGALVPVTAVALCLFVVGSGNVVTVMLIGRYDMRTTNLYLG 78
DB 20 LHKCSNQECHWEEPV-FGMIVCVTIYIPLMLGGLGNILTLVWLRPWRSTYLYLS 78
QY 79 SMAVSDLLILGLPFLVLRWRSRPWFGPLLCRLSLYLGVGCTYATLHMTALSVRYL 138
DB 79 SLAVSDILILLPLVDLYLWLRPWPFGFIFCKSTMPFSECCTFCISILHITFISRYL 138
QY 139 AICRPLRVLVTRRRRRLAIVLWAVALLSAGPFLVLCVE---QDPGIS----- 186
DB 139 AVCWPTAKTVVTRRTTIIGCIWLGAALSAAPVVMVGVSEVQSDQGLSGWRSGAW 199
QY 187 -----VVPGL----- 191
DB 199 TGKEGKGFIIIGRERENDKGLKDSLEBMWMDKEMNECGDNGVTGFKGDKSLEV 258
QY 192 -NGTARIASSPLASSPLMSRAPPPSPGPE---TAAALFGRERCPSP-PAQLGA 244
DB 259 GECTKEQHEGARAAGEEAQNMKEDEGGGGGGTDDGGGRMQVDTRECRCTDYAVSSGL 318
QY 245 LRVLWVTTAYFFLPFLCLSLYLGLIGRELWSSRRPLRGAASGRGRGRHQTQVRLVLYV 304
DB 319 LSAMLVLSNMVFLVPLVLCILGLVLYGLRTLW-----LRS-QISRRDNNNTVMGLGV 372
QY 305 LAFIICWLPFHVGRIIY-----INTE-----DSR----- 328
DB 373 LVFLCWLFPFHVGRTIFFPSLGSDRFCVNAHALDSRVLPLELPPPGALGESDEAAGDAF 432
QY 329 ----- 328
DB 433 SEATGRGDGRAVLDTRGTREARSDEGAAGSPTEATTADPYDAENSTPLDDTHSHSQYP 492
QY 329 MMVFSQYFNIVALQLFYLASINPILYNLSKKYRAAFKLLAR 373
DB 493 LYLVSQYFNVLSSVLFYLSAAINPILYNLSMSRYRHAVHSLPRR 537

RESULT 6

Q8NE20 PRELIMINARY; PRT; 426 AA.
ID Q8NE20
AC Q8NE20;


```

ID Q91Z76 PRELIMINARY; PRT; 395 AA.
AC Q91Z76;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Neuromedin U receptor type 2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Funes S., Hedrick J.A., Yang S., Shan L., Bayne M., Monsma F.J. Jr.,
RA Gustafson E.L.;
RA "Characterization of murine neuromedin U R2 receptor.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AY057384; AAL26695.1; -.
DR InterPro; IPR000276; GPCR Rhodopsin.
DR InterPro; IPR005390; NeuromedinU.
DR InterPro; IPR005392; NeuromedinU2.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PRO0237; GPCR RHODOPSIN.
DR PRINTS; PRO1565; NEUROMEDINUR.
DR PRINTS; PRO1567; NEUROMEDINUR.
DR PROSITE; PS00237; G PROTEIN RECP F1_1; 1.
DR PROSITE; PS00262; G PROTEIN RECP F1_2; 1.
KW G-protein coupled receptor; Receptor; Transmembrane.
SQ SEQUENCE 395 AA; 44844 MW; 0BB540024566903D CRC64;

Query Match 22.2%; Score 478.5; DB 11; Length 395;
Best Local Similarity 31.5%; Pred. No. 1.5e-31;
Matches 115; Conservative 79; Mismatches 112; Indels 59; Gaps 10;

QY 30 CSPFPLGALVPVTAVCCLFVGVSGNVVTVMLIGRYDMRTTNNLYLGSMAVSDLLI-L 88
DB 29 CGPKRSDLSLPVSVVYALIFVVGIGNLLVCLVIARHQTLLKPTNYLFLSLAVSDLLVLL 88
QY 89 LGLPFDLYRLWRSPWVPGELLCRLSLYVGECTYATLLHMTALSVERYLAICRPLRARV 148
DB 89 LGMPLEYELHNYVFLFPGVGCYFKTALFETVCFASILSVTSIERYVAIVHPFRKL 148
QY 149 LVTRRRVRLAIVLWAVALLSAGPFLVLGV--EODPGISVWVGLNGTARIASPLASSP 206
DB 149 ESTRRRLRLSLVWSVSVFSLPNTSINGIKFQFPNGSVFG-----SATCTVTK 200
QY 207 PLMSRAPPPSPGPGTAAALFSRECRPSAQLGALRVMLWVTYAYFF-LPFLCLSI 265
DB 201 PMV-----YNIQATSFLLPMTLSV 226
QY 266 LYGLIGRELSSRRPLRGPAASGRGRHRTV-RVLLVVLAFIICWLPFHVGRIIYINT 324
DB 227 LYLMGLRLKRDESLEADKVTNTHRPSKSVTKMLFVLVFAICWTPPHVDRLPFSF 285
QY 325 EDSRMVFSQYFN---IVALQLFYLASINPILNLSKRYAAAFKLLA-----RK 374
DB 286 VDEWTESLAAVNLHVHVGCVFFYLSSAVNPILNLSRRFR-AAFRNVVSPCKWCHPQ 344
QY 375 SRPRG 379
DB 345 HRPOG 349

RESULT 9
Q8BZ39 PRELIMINARY; PRT; 395 AA.
ID Q8BZ39
AC Q8BZ39;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Neuromedin U receptor type 2.

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Funes S., Hedrick J.A., Yang S., Shan L., Bayne M., Monsma F.J. Jr.,
RA Gustafson E.L.;
RA "Characterization of murine neuromedin U R2 receptor.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AY057384; AAL26695.1; -.
DR InterPro; IPR000276; GPCR Rhodopsin.
DR InterPro; IPR005390; NeuromedinU.
DR InterPro; IPR005392; NeuromedinU2.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PRO0237; GPCR RHODOPSIN.
DR PRINTS; PRO1565; NEUROMEDINUR.
DR PRINTS; PRO1567; NEUROMEDINUR.
DR PROSITE; PS00237; G PROTEIN RECP F1_1; 1.
DR PROSITE; PS00262; G PROTEIN RECP F1_2; 1.
KW G-protein coupled receptor; Receptor; Transmembrane.
SQ SEQUENCE 395 AA; 44844 MW; 0BB540024566903D CRC64;

Query Match 22.2%; Score 478.5; DB 11; Length 395;
Best Local Similarity 31.5%; Pred. No. 1.5e-31;
Matches 115; Conservative 79; Mismatches 112; Indels 59; Gaps 10;

QY 30 CSPFPLGALVPVTAVCCLFVGVSGNVVTVMLIGRYDMRTTNNLYLGSMAVSDLLI-L 88
DB 29 CGPKRSDLSLPVSVVYALIFVVGIGNLLVCLVIARHQTLLKPTNYLFLSLAVSDLLVLL 88
QY 89 LGLPFDLYRLWRSPWVPGELLCRLSLYVGECTYATLLHMTALSVERYLAICRPLRARV 148
DB 89 LGMPLEYELHNYVFLFPGVGCYFKTALFETVCFASILSVTSIERYVAIVHPFRKL 148
QY 149 LVTRRRVRLAIVLWAVALLSAGPFLVLGV--EODPGISVWVGLNGTARIASPLASSP 206
DB 149 ESTRRRLRLSLVWSVSVFSLPNTSINGIKFQFPNGSVFG-----SATCTVTK 200
QY 207 PLMSRAPPPSPGPGTAAALFSRECRPSAQLGALRVMLWVTYAYFF-LPFLCLSI 265
DB 201 PMV-----YNIQATSFLLPMTLSV 226
QY 266 LYGLIGRELSSRRPLRGPAASGRGRHRTV-RVLLVVLAFIICWLPFHVGRIIYINT 324
DB 227 LYLMGLRLKRDESLEADKVTNTHRPSKSVTKMLFVLVFAICWTPPHVDRLPFSF 285
QY 325 EDSRMVFSQYFN---IVALQLFYLASINPILNLSKRYAAAFKLLA-----RK 374
DB 286 VDEWTESLAAVNLHVHVGCVFFYLSSAVNPILNLSRRFR-AAFRNVVSPCKWCHPQ 344
QY 375 SRPRG 379
DB 345 HRPOG 349

RESULT 10
Q9JIB1 PRELIMINARY; PRT; 395 AA.
ID Q9JIB1
AC Q9JIB1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Neuromedin U receptor 2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=20351041; PubMed=10894543;
RA Howard A.D., Wang R., Pong S.-S., Mellin T.N., Strack A., Guan X.-M.,
RA Zeng Z., Williams D.L., Feighner S.D., Nunes C.N., Murphy B.,
RA Stair J.N., Yu H., Jiang Q., Clements M.K., Tan C.P., McKee K.K.,
RA Hreniuk D.L., McDonald T.P., Lynch K.R., Evans J.P., Austin C.P.,
RA Caskey T., van der Ploeg L.H.T., Liu Q.;
RT "Identification of receptors for neuromedin U and its role in feeding.";

```


the Central Nervous System."; J. Biol. Chem. 275:39482-39486 (2000).
 [3]
 RY SEQUENCE FROM N.A.
 RA Pang L., Wang S., Laz T., Hedrick J.A.;
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 [4]
 RP SEQUENCE OF 4-415 FROM N.A.
 RX MEDLINE=20351041; PubMed=10894543;
 RA Howard A.D., Wang R., Pong S.-S., Mellin T.N., Strack A., Guan X.-M.,
 RA Zeng Z., Williams D.L., Feighner S.D., Nunes C.N., Murphy B.,
 RA Stair J.N., Yu H., Jiang Q., Clemente M.K., Tan C.P., McKee K.K.,
 RA Hreniuk D.L., McDonald T.P., Lynch K.R., Evans J.F., Austen C.P.,
 RA Caskey T., van der Ploeg L.H.T., Liu Q.;
 RT "Identification of receptors for neuromedin U and its role in
 RT feeding."; Nature 406:70-74 (2000).
 RL [5]
 RP SEQUENCE OF 4-415 FROM N.A.
 RX PubMed=10887190;
 RA Hosoya M., Moriya T., Kawamata Y., Ohkubo S., Fujii R., Matsui H.,
 RA Shintani Y., Fukusumi S., Habata Y., Hinuma S., Onda H., Nishimura O.,
 RA Fujino M.;
 RT "Identification and Functional Characterization of a Novel Subtype of
 RT Neuromedin U Receptor."; J. Biol. Chem. 275:29528-29532 (2000).
 RL J. Biol. Chem. 275:29528-29532 (2000).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 DR EMBL: AF272363; AAG24794.1; -;
 DR EMBL: AF292402; AAG03064.1; -;
 DR EMBL: AF242874; AAF82755.1; -;
 DR EMBL: AB041228; BAB13721.1; -;
 DR InterPro: IPR005390; NeuroMedinU.
 DR InterPro: IPR000276; GPCR_Rhodopsin.
 DR InterPro: IPR005392; NeuroMedinU2.
 DR Pfam: PF00001; 7tm 1; 1.
 DR PRINTS: PR00237; GPCRRHODOPSIN.
 DR PRINTS: PR01565; NEUROMEDINUR.
 DR PRINTS: PR01567; NEUROMEDINUR.
 DR PROSITE: PS00237; G-PROTEIN RECEPTOR F1.1; 1.
 DR PROSITE: PS0262; G-PROTEIN RECEPTOR F1.2; 1.
 DR PROSITE: PS0262; G-PROTEIN RECEPTOR F1.2; 1.
 KW G-protein coupled receptor; Receptor; Transmembrane.
 SQ SEQUENCE 415 AA; 47725 MW; C2BACD84B313390F CRC64;
 Query Match 22.1%; Score 476; DB 4; Length 415;
 Best Local Similarity 31.8%; Pred. No. 2.66-31;
 Matches 115; Conservative 81; Mismatches 106; Indels 60; Gaps 11;
 QY 30 CSPFPLGALVPVTAVALCLFVGVGNVTVMLIGRYDMRTTNNLYLGSMAVSDLLI-L 88
 DB 37 CGPRSHFFLPVSVVYVPIFVGVGNVTVMLIGRYDMRTTNNLYLGSMAVSDLLI-L 96
 QY 89 LGLPFDLYRLWRSRPVWFGPFLCLRLSLYVGECCGYATLLHMTALSVERYLAICRPLR 148
 DB 97 LGMPLLEYEMWNNYVPLFGVGVGNVTVMLIGRYDMRTTNNLYLGSMAVSDLLI-L 156
 QY 149 LVTRRRRLALIAVLAVALLSAGPFLVGV--QDPGISVWPGINGTARIASPLASSP 206
 DB 157 QSTERRALRLIGVWGSFVLSPLNTSINGIKFHYFNGSLVFG-----SATCTVIK 208
 QY 207 PLWSRAPPPSPSGPETAEAAALFSRECRPSAQLGALRVLMLVTTAYFF-LPFLCLSI 265
 DB 209 PMWI-----YNFIIQVTSFLLPMTVISV 234
 QY 266 LYGLIGRELWSSRRPLRGPAASGRGRHRTV-RVLLVVVLAFLIICWLPFHVGRIIYINT 324
 DB 235 LYLLMALRLKKDSLEADGNANTQRCRKSVMKLFVLVLAIFAICWAPPHIDLFPFV 294
 QY 325 ED--SRMYFSQYFNIVAL---QLFYLASINPILYNLISKYRAAFKLLARKSRPRGF 380
 DB 295 EEWSESL--RAVFNLVHVSVGVFFYLSAVNPILYNLISRRFQ-AAFQNVIS-----SP 345
 QY 381 HR 382

Db 346 HK 347

RESULT 13

Q9ESQ4 PRELIMINARY; PRT; 395 AA.
 AC Q9ESQ4;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE G-protein-coupled receptor TGR-1.
 GN TGR-1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20449029; PubMed=10887190;
 RA Hosoya M., Moriya T., Kawamata Y., Ohkubo S., Fujii R., Matsui H.,
 RA Shintani Y., Fukusumi S., Habata Y., Hinuma S., Onda H., Nishimura O.,
 RA Fujino M.;
 RT "Identification and Functional Characterization of a Novel Subtype of
 RT Neuromedin U Receptor."; J. Biol. Chem. 275:29528-29532 (2000).
 RL J. Biol. Chem. 275:29528-29532 (2000).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 DR EMBL: AB041229; BAB13722.1; -;
 DR InterPro: IPR000276; GPCR_Rhodopsin.
 DR InterPro: IPR005390; NeuroMedinU.
 DR InterPro: IPR005392; NeuroMedinU2.
 DR Pfam: PF00001; 7tm 1; 1.
 DR PRINTS: PR00237; GPCRRHODOPSIN.
 DR PRINTS: PR01565; NEUROMEDINUR.
 DR PRINTS: PR01567; NEUROMEDINUR.
 DR PROSITE: PS00237; G-PROTEIN RECEPTOR F1.1; 1.
 DR PROSITE: PS0262; G-PROTEIN RECEPTOR F1.2; 1.
 DR PROSITE: PS0262; G-PROTEIN RECEPTOR F1.2; 1.
 KW G-protein coupled receptor; Receptor; Transmembrane.
 SQ SEQUENCE 395 AA; 44756 MW; EABD6D36E9C355DA CRC64;
 Query Match 22.1%; Score 475.5; DB 11; Length 395;
 Best Local Similarity 31.5%; Pred. No. 2.7e-31;
 Matches 108; Conservative 75; Mismatches 111; Indels 49; Gaps 7;
 QY 30 CSPFPLGALVPVTAVALCLFVGVGNVTVMLIGRYDMRTTNNLYLGSMAVSDLLI-L 88
 DB 29 CGPRSDLSLPVSAVALFLVGVGNVTVMLIGRYDMRTTNNLYLGSMAVSDLLI-L 88
 QY 89 LGLPFDLYRLWRSRPVWFGPFLCLRLSLYVGECCGYATLLHMTALSVERYLAICRPLR 148
 DB 89 LGMPLLEYEMWNNYVPLFGVGVGNVTVMLIGRYDMRTTNNLYLGSMAVSDLLI-L 148
 QY 149 LVTRRRRLALIAVLAVALLSAGPFLVGV--EQDPGISVWPGINGTARIASPLASSP 206
 DB 149 ESTRRRALRLISLVWGSFVLSPLNTSINGIKFQHPFNGSSVFG-----SATCTVTK 200
 QY 207 PLWSRAPPPSPSGPETAEAAALFSRECRPSAQLGALRVLMLVTTAYFF-LPFLCLSI 265
 DB 201 PMWV-----YNLIQVTSFLLPMTVISV 226
 QY 266 LYGLIGRELWSSRRPLRGPAASGRGRHRTV-RVLLVVVLAFLIICWLPFHVGRIIYINT 324
 DB 227 LYLLMALRLKKDSLEANKVAVNIHPSKSVTKMLFVLVLAIFAICWTPPHVDRLPFV 286
 QY 325 ED--SRMYFSQYFNIVALQOLFYLASINPILYNLISKYRAA 365
 DB 287 EEWTESLAAVFNLIHVSVGVFFYLSAVNPILYNLISRRPRAA 329
 RESULT 14
 O55040
 ID O55040 PRELIMINARY; PRT; 405 AA.

OS05040;
 01-JUN-1998 (TrEMBLrel. 06, Created)
 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 01-OCT-2002 (TrEMBLrel. 23, Last annotation update)
 Orphan G protein-coupled receptor.
 GPR66.
 Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA McKee K.K., Tan C.T., Liu J., Palyha O.C., Feighner S.D.,
 RA Hreniuk D.L., Smith R.G., Howard A.D.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 DR EMBL; AF044602; AAC02681.1; -;
 DR MGD; MG1:1341898; Gpr66.
 DR InterPro; IPR000276; GPCR Rhodopsin.
 DR InterPro; IPR005390; NeuromedinU.
 DR InterPro; IPR005391; NeuromedinU.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PRO0237; GPCR RHODOPSIN.
 DR PRINTS; PRO1565; NEUROMEDINUR.
 DR PRINTS; PRO1566; NEUROMEDINUR.
 DR PROSITE; PS00237; G PROTEIN RECEPTOR FL 1; 1.
 DR PROSITE; PS0262; G PROTEIN RECEPTOR FL 2; 1.
 DR G-protein coupled receptor; Receptor; Transmembrane.
 KW G-protein coupled receptor; Receptor; Transmembrane.
 SQ SEQUENCE 405 AA; 45609 MW; F1BA493D3EB81F34 CRC64;
 Query Match 21.2%; Score 457.5; DB 11; Length 405;
 Best Local Similarity 32.2%; Pred. No. 8.6e-30;
 Matches 133; Conservative 62; Mismatches 141; Indels 77; Gaps 14;
 QY 19 WPALPP-----CDE-----RRCSPPFLGALVPTAVCLCLFVVGSGNNVVTVMILGRVDMR 70
 DB 10 WPKQEDNLNLTDEALRLKYLGPQOMKQFVPCVYLLIFVVGTLGNGLTCTVILRNKTM 69
 QY 71 TTNLYLGSMAVSDLLILL-GLPDLVRLRSRPPWVFGPDLCLRLSLVVGSGCTYATLLHM 129
 DB 70 TPTNLYLGSMAVSDLLILL-GLPDLVRLRSRPPWVFGPDLCLRLSLVVGSGCTYATLLHM 129
 QY 130 TALSVRYLAICRPLRARLVTRRRVRALIAVMAVALLSAGPFLFVVGVEQPGISVWP 189
 DB 130 TALSVRYLAICRPLRARLVTRRRVRALIAVMAVALLSAGPFLFVVGVEQPGISVWP 189
 QY 130 TALSVRYLAICRPLRARLVTRRRVRALIAVMAVALLSAGPFLFVVGVEQPGISVWP 189
 DB 130 TALSVRYLAICRPLRARLVTRRRVRALIAVMAVALLSAGPFLFVVGVEQPGISVWP 189
 QY 190 GLNGTARIASSPLASSPPLMWSRAPPPSPGPETAFAAALFGRCPSPALGALRVML 249
 DB 186 -----CRGFVDP-----SAICS-----LVGPMDFVKL-VVL 210
 QY 250 WVTAYEFLPCLLSILYGLIGRELWSSRRPL-----RGPAASG-----RRGHR 294
 DB 211 TTALLFFCLPMVTVISVLYLIGLRLRRERMLLQVEVKGKTAQTQETSHRRIQLQDRGR 270
 QY 295 QTVRVLLVAVLAFICMLPFHVGRIIYI-----NTDSRMWYFSQYFNIVALQFLVLSAS 349
 DB 271 QVTQMLFALVVVGICWAPHADRMWSLVYGHSTEGHLAY--QCWHIASGIFFFYLGSA 328
 QY 350 INPILYNLISKYRAAFAKLL-----LARKSRPRGPHRS-RDTAGVAGDTG 395
 DB 329 ANPVLYSLMSTRFETFLQALGLGTQCCHRRQPVGHSHNHLRLTGTSLCDVG 381

RESULT 15

Q8ITC7

ID Q8ITC7

AC Q8ITC7

DT 01-MAR-2003

DT 01-MAR-2003

DT 01-MAR-2003

DE Putative Cap2b receptor.

GN CG14575.

PRELIMINARY;

PRT; 477 AA.

(TrEMBLrel. 23, Created)

(TrEMBLrel. 23, Last sequence update)

(TrEMBLrel. 23, Last annotation update)

OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Park Y., Kim Y.J., Adams M.E.;
 RA MEDLINE=22177201; PubMed=12177421;
 RT "Identification of G protein-coupled receptors for Drosophila PRXamide
 RT peptides, CCAP, corazonin, and AKH supports a theory of ligand-
 RT receptor coevolution."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:11423-11428(2002).
 DR EMBL; AF522193; AAN10046.1; -;
 KW Receptor.
 SQ SEQUENCE 477 AA; 54004 MW; 34FFB13F301A9F4D CRC64;
Query Match 21.0%; Score 453; DB 5; Length 477;
 Best Local Similarity 29.7%; Pred. No. 2.4e-29;
 Matches 117; Conservative 68; Mismatches 135; Indels 74; Gaps 11;
QY 8 SDGEGAREPPWALPPCDERRCSP-----FPLGALVPTAVCLCLFVVGSG 55
 DB 30 SDPSHGFGEEDY-----ACGTFCNSKPEFVAVLGPQTLPLYKAVLITIIIFGGIFITGVG 85
 QY 56 NVVTVMILGRYDMRTTNLYLGSMAVSDLLILL-GLPDLVRLRSRPPWVFGPDLCLRLS 114
 DB 86 NLLVCIVILRHSAMHTATNYLFSLAVSDLLYLLFGLPTEVFLYWHQYDPLFGMPFCKIR 145
 QY 115 LVVGECCTYATLLHMTALSVERYLAI CRPLRARLVTRRRVRALIAVMAVALLSAGPFL 174
 DB 146 AFISEACTVSVFTIVAFSMERFLAICHPLHLVAMVGFKAIRIITALTIVTSISAIPP- 204
 QY 175 FLVGVEQDPOISVVGPGINGTARIASSPLASSPPLMWSRAPPPSPGPETAFAAALFSRE 234
 DB 205 -----GLSDIQYINYPDHS-----RIEESAF----- 227
 QY 235 CRPSPAQALGALRVMLVTTAYRPLPFLCLLSILYGLIGRELWSSRRPLRG--PAASGRERG 292
 DB 228 CSMSPKIVNEIPVFEVSCIFVFPIMLILLYGRMGAKIRSTNOKLGVQOQTNNRETR 287
 QY 293 HRQ-----TVRVLLVVLAFIICWLPFHVGRIIYIINTDSRMWYFSQYFNI-----VAL 341
 DB 288 NSQMRKKTIVIRLAAVITFFVCWCFPHLQRLIFLVAKN-----MDNYLDINEALFSIAG 342
 QY 342 QLPYLSASINPILYNLISKYRAAFAKLLARKS 375
 DB 343 FAYTVSCTVNPVIVYSVMRRYR-VAPRELICGKA 375
Search completed: January 1, 2004, 06:22:26
 Job time : 71.6341 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 31, 2003, 18:55:58 ; Search time 5295.11 Seconds
(without alignments)
10739.032 Million cell updates/sec

Title: US-09-719-485-4
Perfect score: 1390
Sequence: 1 atggggcagcccttggaacgg.....acgtgaagacgatgggataa 1390

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1161	83.5	1161	6	AX154591	Sequence AX154591
2	1078	77.6	1239	6	AX154589	Sequence AX154589
3	1078	77.6	1239	6	AX549187	Sequence AX549187
4	1078	77.6	1239	6	AX572965	Sequence AX572965
5	1052	75.7	2040	9	AF034632	Homo sapi AF034632
6	1052	75.7	163284	9	AL137000	Human DNA AL137000
7	1052	75.7	341560	2	AL596304	Homo sapi AL596304
8	1052	75.7	349980	6	AX711879	Sequence AX711879
9	1052	75.7	349980	6	AX739961	Sequence AX739961
10	592.2	42.6	692	9	HS3339407	Homo sapi HS3339407
11	572.2	41.2	692	9	HS3342408	Homo sapi HS3342408
12	557.4	40.1	615	9	HS3339459	Homo sapi HS3339459
13	424.8	30.6	701	9	HS3326768	Homo sapi HS3326768
14	307	22.1	192116	2	AC109158	Mus muscu AC109158
15	283	20.4	283	6	AR168467	Sequence AR168467
16	283	20.4	283	6	AR182284	Sequence AR182284
17	283	20.4	283	6	AR303910	Sequence AR303910
18	283	20.4	283	6	BD056681	Human gal BD056681
19	283	20.4	283	6	BD064652	Human gal BD064652
20	283	20.4	283	6	BD064664	Mouse gal BD064664
21	266.8	19.2	573	9	HS3341295	Homo sapi HS3341295
22	242.6	17.5	1351	5	AB095997	Gallus ga AB095997
23	242.6	17.5	1703	5	AB095996	Gallus ga AB095996
24	242.6	17.5	1751	5	AB095995	Gallus ga AB095995
25	242.6	17.5	4121	5	AB095994	Gallus ga AB095994
26	234.8	16.9	1063	6	AR156351	Sequence AR156351
27	234.8	16.9	1101	4	SSU60178	Sub scrofa SSU60178
28	233.2	16.8	870	4	SSU60180	Sub scrofa SSU60180
29	230	16.5	1029	6	AR156352	Sequence AR156352
30	230	16.5	110000	2	AC095360_2	Continuation (3 of AC117349 Rattus no AC117349)
31	230	16.5	183099	2	AC117349	Rattus no AC117349
32	230	16.5	195131	2	AC108400	Mus muscu AC108400
33	230	16.5	227035	2	AC116585	Mus muscu AC116585
34	230	16.5	228796	2	AC095173	Rattus no AC095173
35	230	16.5	250754	2	AC130771	Rattus no AC130771
36	229	16.5	250	6	AR269796	Sequence AR269796
37	228.8	16.5	870	6	AX548854	Sequence AX548854
38	228.8	16.5	870	9	HSU60181	Human growt HSU60181
39	228.8	16.5	1122	6	AR156354	Sequence AR156354
40	228.8	16.5	6787	9	AF369786	Homo sapi AF369786
41	228.8	16.5	145593	9	AC069523	Homo sapi AC069523
42	228.8	16.5	159005	2	AC015655	Homo sapi AC015655
43	228.8	16.4	1092	6	AR156357	Sequence AR156357
44	228.4	16.4	1092	6	BD094703	GHSR liga BD094703
45	228.4	16.4	1092	6	E54510	Novel polyp E54510

ALIGNMENTS

RESULT 1	AX154591	AX154591	1161 bp	DNA	linear	PAT 23-JUN-2001
LOCUS	Sequence 11	from Patent WO0138355.				
DEFINITION	Sequence 11	from Patent WO0138355.				
ACCESSION	AX154591					
VERSION	AX154591.1	GI:14536177				
KEYWORDS	Homo sapiens (human)					
SOURCE	Homo sapiens					
ORGANISM	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1	Sheppard, P.O., Jaspers, S.R., Deisher, T.A. and Bishop, P.D.				
AUTHORS		Method of forming a peptide-receptor complex with zaig33 and				
TITLE		therapeutic use thereof				

Best Local Similarity 89.1%; Pred. No. 4.5e-134;
Matches 1239; Conservative 0; Mismatches 0; Indels 151; Gaps 1;

QY 1 ATGGGAGAGCCCTTGGAAACGACGACGCGCCCGAGGGGCGCGGGAGCCCGTGGGCC 60
DB 1 ATGGGAGAGCCCTTGGAAACGACGACGCGCCCGAGGGGCGCGGGAGCCCGTGGGCC 60
QY 61 GCCTGCGCGCTTGGAGAGCGCGCTGTCTGCGCCCTTTCCCTGGGGGCGCTGGTGGCG 120
DB 61 GCCTGCGCGCTTGGAGAGCGCGCTGTCTGCGCCCTTTCCCTGGGGGCGCTGGTGGCG 120
QY 121 GTGACCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 180
DB 121 GTGACCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 180
QY 181 ATGCTGATCGGGCGCTACCGGGACATCGCGACCAACCACTTGTACCTGGGCGAGCATG 240
DB 181 ATGCTGATCGGGCGCTACCGGGACATCGCGACCAACCACTTGTACCTGGGCGAGCATG 240
QY 241 GCCGTGTCGACCTTACTCATCTGCTCGGGCTGCGGTGACCTGTACCGCTCTGGCGC 300
DB 241 GCCGTGTCGACCTTACTCATCTGCTCGGGCTGCGGTGACCTGTACCGCTCTGGCGC 300
QY 301 TCGGGCCCTGGGTGTGCGGGCGCGCTGTCTGCGCGCTGTCTGCGCGCTGTCTGCGCGCT 360
DB 301 TCGGGCCCTGGGTGTGCGGGCGCGCTGTCTGCGCGCTGTCTGCGCGCTGTCTGCGCGCT 360
QY 361 TGCACCTTACCGCCAGCTGTGTCATGACGCGCTCAGCGCTCAGCGCTCAGCGCTACCTGGCCATC 420
DB 361 TGCACCTTACCGCCAGCTGTGTCATGACGCGCTCAGCGCTCAGCGCTCAGCGCTACCTGGCCATC 420
QY 421 TGGCGCCCGCTCGCGCGCCCGCTTGTGTCACCGCGCGCGCTGTCGCGCGCTCATCGCT 480
DB 421 TGGCGCCCGCTCGCGCGCCCGCTTGTGTCACCGCGCGCGCTGTCGCGCGCTCATCGCT 480
QY 481 GTGCTGTGGCGCTGGCGCTGTCTGCGCGCTCCCTTCTGCTGTGCTGTGCTGTGCTGTGCT 540
DB 481 GTGCTGTGGCGCTGGCGCTGTCTGCGCGCTCCCTTCTGCTGTGCTGTGCTGTGCTGTGCT 540
QY 541 CAGGACCCCGGCTATCCGTAGTCCCGGGCTCAATGGCACCGCGCGGATCGCTCTCTCG 600
DB 541 CAGGACCCCGGCTATCCGTAGTCCCGGGCTCAATGGCACCGCGCGGATCGCTCTCTCG 600
QY 601 CCTCTCGCTCTGCGCGCTCTCTGCTGTCTGCGGGCGCACCGCGCTGCGCGCGCTCG 660
DB 601 CCTCTCGCTCTGCGCGCTCTCTGCTGTCTGCGGGCGCACCGCGCTGCGCGCGCTCG 660
QY 661 GGGCCCGAGACCGCGAGCGCGCGCTGTTCAGCGCGGAAATGCTGCGCGAGCCCGCGCG 720
DB 661 GGGCCCGAGACCGCGAGCGCGCGCTGTTCAGCGCGGAAATGCTGCGCGAGCCCGCGCG 720
QY 721 CAGCTGGCGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 780
DB 721 CAGCTGGCGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 780
QY 781 CTGTGCTCTAGACCTCTACGGGCTCATCGGGGAGAGCTGTGAGCAGCGCGCGCGCG 840
DB 781 CTGTGCTCTAGACCTCTACGGGCTCATCGGGGAGAGCTGTGAGCAGCGCGCGCGCG 840
QY 841 CTGCGAGCGCGCGCTCTGGGGCGGAGAGGCCACCGCGAGACCGTCCGCGTCTG 900
DB 841 CTGCGAGCGCGCGCTCTGGGGCGGAGAGGCCACCGCGAGACCGTCCGCGTCTG 900
QY 901 CGTAAGTGGAGCGCGCTGTGCTTCAAAGAGCGCTGTGCTGTGCTGTGCTGTGCTGTGCT 960
DB 901 C----- 901
QY 961 GCGCAACGCTGGGTCTCCCTTCTCCCTGTCTGCGCGCTCTGGGCGCGCTTCCAGCTCCC 1020
DB 902 ----- 901
QY 1021 TTTCCTATTGATTCAGGCTTCCACCGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1080
DB 1021 TTTCCTATTGATTCAGGCTTCCACCGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1080

Db 902 -----TGGTGGTGGTCTCGCATTTATAATTG 929
QY 1081 CTGTTGGCCCTTCCAGCTTGGCAGAAATCATTTACATAAAACACGGAAGATTCGCGGATGAT 1140
DB 930 CTGTTGGCCCTTCCAGCTTGGCAGAAATCATTTACATAAAACACGGAAGATTCGCGGATGAT 989
QY 1141 GTACTTCTCTCAGTACTTTAAACATCGCTCGCTGTGCAACTTTTCTATCTGAGCGCATCTAT 1200
DB 990 GTACTTCTCTCAGTACTTTAAACATCGCTCGCTGTGCAACTTTTCTATCTGAGCGCATCTAT 1049
QY 1201 CAACCAATCTCTCAACCTCATTTCAAGAGTACAGAGCGCGCCCTTTAAACTGCT 1260
DB 1050 CAACCAATCTCTCAACCTCATTTCAAGAGTACAGAGCGCGCCCTTTAAACTGCT 1109
QY 1261 GCTCGCAAGAGTCCAGCGCGAGAGCTTCCACAGAAGCAGGACACATGCGGGGGAAGT 1320
DB 1110 GCTCGCAAGAGTCCAGCGCGAGAGCTTCCACAGAAGCAGGACACATGCGGGGGAAGT 1169
QY 1321 TGCAGGGGACATGAGAGGAGACACGCTGGGCTACACCGAGACAAGCGCTAACGTGAAGAC 1380
DB 1170 TGCAGGGGACATGAGAGGAGACACGCTGGGCTACACCGAGACAAGCGCTAACGTGAAGAC 1229
QY 1381 GATGGGATAA 1390
DB 1230 GATGGGATAA 1239

RESULT 3
AX549187 LOCUS 1239 bp DNA linear PAT 26-NOV-2002
DEFINITION Sequence 472 from Patent WO02061087.
ACCESSION AX549187
VERSION AX549187.1 GI:25813904
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Burner, G.C., Roush, C.L. and Brown, J.P.
TITLE Antigenic peptides, such as for G protein-coupled receptors (GPCRs), antibodies thereto, and systems for identifying such antigenic peptides
JOURNAL Patent: WO 02061087-A 472 08-AUG-2002;
Lifespan Biosciences, Inc. (US)
FEATURES Location/Qualifiers
source 1..1239
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 172 a 431 c 397 g 239 t
ORIGIN

Query Match 77.6%; Score 1078; DB 6; Length 1239;
Best Local Similarity 89.1%; Pred. No. 4.5e-134;
Matches 1239; Conservative 0; Mismatches 0; Indels 151; Gaps 1;

QY 1 ATGGGAGAGCCCTTGGAAACGACGACGCGCCCGAGGGGCGCGGGAGCCCGTGGGCC 60
DB 1 ATGGGAGAGCCCTTGGAAACGACGACGCGCCCGAGGGGCGCGGGAGCCCGTGGGCC 60
QY 61 GCCTGCGCGCTTGGAGAGCGCGCTGTCTGCGCCCTTTCCCTGGGGGCGCTGGTGGCG 120
DB 61 GCCTGCGCGCTTGGAGAGCGCGCTGTCTGCGCCCTTTCCCTGGGGGCGCTGGTGGCG 120
QY 121 GTGACCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 180
DB 121 GTGACCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 180
QY 181 ATGCTGATCGGGCGCTACCGGGACATCGCGACCAACCACTTGTACCTGGGCGAGCATG 240
DB 181 ATGCTGATCGGGCGCTACCGGGACATCGCGACCAACCACTTGTACCTGGGCGAGCATG 240

Qy	241	GCCTGTCCGACCTACTCATCTCTGCTCGGGCTGCCGTTGACCTGTACCGCCTCTGCGC	300
Db	241	GCCTGTCCGACCTACTCATCTCTGCTCGGGCTGCCGTTGACCTGTACCGCCTCTGCGC	300
Qy	301	TCGGGGCCCTGGGTGTTTCGGGCCCTGTCTCTGCGCGCTGTCCCTCTACGTGGGGAGGGC	360
Db	301	TCGGGGCCCTGGGTGTTTCGGGGCCCTGTCTCTGCGCGCTGTCCCTCTACGTGGGGAGGGC	360
Qy	361	TGCACTACGCACCGCTGCTGCACATGACCCGCGCTCAGCGCTCGAGCGCTACCTGGCCATC	420
Db	361	TGCACTACGCACCGCTGCTGCACATGACCCGCGCTCAGCGCTCGAGCGCTACCTGGCCATC	420
Qy	421	TGCGCCCGCTCCGCGCCGCGCTTGTGTACACCGCGCGCGCTGTCGCGCGCTCATCGCT	480
Db	421	TGCGCCCGCTCCGCGCCGCGCTTGTGTACACCGCGCGCGCTGTCGCGCGCTCATCGCT	480
Qy	481	GTGCTCTGGGCGGTGGGCGCTCTCTGCGCGCTCCCTTCTGTTCCTGTGGGCGTGCAG	540
Db	481	GTGCTCTGGGCGGTGGGCGCTCTCTGCGCGGTCCCTTCTGTTCCTGTGGGCGTGCAG	540
Qy	541	CAGGACCCCGGCATCTCCGTAGTCCCGGGCTCAATGGCACCGCGCGGATCGCCTCCTCG	600
Db	541	CAGGACCCCGGCATCTCCGTAGTCCCGGGCTCAATGGCACCGCGCGGATCGCCTCCTCG	600
Qy	601	CCTCTCGCTGTGTCGCGCCTCTCTGTGCTCTGCGGGCGCAACCGCGTCCCGCGCTCG	660
Db	601	CCTCTCGCTGTGTCGCGCCTCTCTGTGCTCTGCGGGCGCAACCGCGTCCCGCGCTCG	660
Qy	661	GGGGCCGAGACCGGGAGGCGCGCGCTGTGTAGCCGCGAAATGCGCGCGAGCCCGCG	720
Db	661	GGGGCCGAGACCGGGAGGCGCGCGCTGTGTAGCCGCGAAATGCGCGCGAGCCCGCG	720
Qy	721	CAGCTGGCGCGCTGCGGTGTCATGCTGTGGGTACACACCGCTACTTCTCTGCGCCTTT	780
Db	721	CAGCTGGCGCGCTGCGGTGTCATGCTGTGGGTACACACCGCTACTTCTCTGCGCCTTT	780
Qy	781	CTGTGCTCTAGCATCTCTACGGGCTCATCGGGCGGAGCTGTGGAGCAGCCGCGCGCG	840
Db	781	CTGTGCTCTAGCATCTCTACGGGCTCATCGGGCGGAGCTGTGGAGCAGCCGCGCGCG	840
Qy	841	CTGCGAGGCGCGCGCCTCGGGCGGGAGAGGCCACCGGCAGACGCTCCGGTCTTG	900
Db	841	CTGCGAGGCGCGCGCCTCGGGCGGGAGAGGCCACCGGCAGACGCTCCGGTCTTG	900
Qy	901	CGTAAGTGAGAGCCCGTGTTCAAAGA CGCTCTGCACTCCGCCCGCGCGGGAGC	960
Db	901	C-----	901
Qy	961	GCGCAAAACGTGGTGCCCTTCCCTCTGCTCGCCAGCTCTGGGCGCGCTTCCAGTCCC	1020
Db	902	-----	901
Qy	1021	TTTCTATTTGATTCCAGCCTCACCCGCGTGTGTGGTCTTGCGATTTTAAATTG	1080
Db	902	-----TGTTGGTGGTCTTGCGATTTTAAATTG	929
Qy	1081	CTGGTTGCCCTTCCACGTTGSCAGATCATTTACATAAACAAGAACATTCGGGATGAT	1140
Db	930	CTGGTTGCCCTTCCACGTTGSCAGATCATTTACATAAACAAGAACATTCGGGATGAT	989
Qy	1141	GTACTTCTCAGTACTTTAAACATCGTCTGCAACTTTCTATCTGAGCGCATCTAT	1200
Db	990	GTACTTCTCAGTACTTTAAACATCGTCTGCAACTTTCTATCTGAGCGCATCTAT	1049
Qy	1201	CAACCCAACTCTTACAACTCATTTCAAAGAAGTACAGAGCGCGCCTTTAACTGCT	1260
Db	1050	CAACCCAACTCTTACAACTCATTTCAAAGAAGTACAGAGCGCGCCTTTAACTGCT	1109
Qy	1261	GCTCGAAGGAAGTCCAGGCGGAGAGCTTCCACAGAGCAGGAGCACTCGGGGGAAGT	1320
Db	1110	GCTCGAAGGAAGTCCAGGCGGAGAGCTTCCACAGAAGCAGGAGCACTCGGGGGAAGT	1169
Qy	1321	TGCAGGGGACACTGGAGGAGACAGGTGGGCTACCGGAGACAGCGCTAACGTGAAGAC	1380

Db	1170	TGCAGGGGACACTGGAGGAGACACGGTGGGCTACACCGAGACAAGCGCTAACGTGAAGAC	1239
Qy	1381	GATGGGATAA	1390
Db	1230	GATGGGATAA	1239
RESULT 4			
AX572965			
LOCUS	AX572965	1239 bp	DNA
DEFINITION	Sequence 1 from Patent WO02057791.		linear
ACCESSION	AX572965		
VERSION	AX572965.1	GI:26005013	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
REFERENCE			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
JOURNAL	Brown, J.P., Burmer, G.C., Roush, C.L. and Kulander, B.G.		
FEATURES	Diagnostic and therapeutic compositions and methods related to gpr		
source	1. .1239		
BASE COUNT	172 a	431 c	397 g
ORIGIN		239 t	
Query Match	77.6%	Score 1078;	DB 6; Length 1239;
Best Local Similarity	89.1%;	Pred. No. 4.5e-134;	
Matches 1239;	Conservative	0; Mismatches	0; Indels 151; Gaps
Qy	1	ATGGGGAGCCCTGGGAACGGCAGCGACGGCCCGAGGGGGCGGGAGCGCGGTGGCCCC	60
Db	1	ATGGGGAGCCCTGGGAACGGCAGCGACGGCCCGAGGGGGCGGGAGCGCGGTGGCCCC	60
Qy	61	CGCGTGGCGGCTTGGCAGACGAGCGCGGCTGCTCGGCCCTTTCCCTTGGGGGGCGCTGGTGGCCG	120
Db	61	CGCGTGGCGGCTTGGCAGACGAGCGCGGCTGCTCGGCCCTTTCCCTTGGGGGGCGCTGGTGGCCG	120
Qy	121	GTGACCGCTGTGTGCTGTGTCCTTTCGTTCGTTCGTTCGGGTGAGCGGCAACGTTGTGACCGGT	180
Db	121	GTGACCGCTGTGTGCTGTGTCCTTTCGTTCGTTCGTTCGGGTGAGCGGCAACGTTGTGACCGGT	180
Qy	181	ATGCTGATCGGGCGCTACCGGGACATCGGACACACCAACCTGTACCTGGGCGACGATG	240
Db	181	ATGCTGATCGGGCGCTACCGGGACATCGGACACACCAACCTGTACCTGGGCGACGATG	240
Qy	241	GCGGTGTCCGACCTACTCATCTTCTCGGGCTGCGGCTTCGACCTGTACCGCCCTCTGGCGC	300
Db	241	GCGGTGTCCGACCTACTCATCTTCTCGGGCTGCGGCTTCGACCTGTACCGCCCTCTGGCGC	300
Qy	301	TCGGGGCCCTGGGGTTCGGGGCGGCTGTCTGTGCGGCTGTTCCTCTACGTGGGGGAGGGC	360
Db	301	TCGGGGCCCTGGGGTTCGGGGCGGCTGTCTGTGCGGCTGTTCCTCTACGTGGGGGAGGGC	360
Qy	361	TGCACCTACCGCAGCTGTGCATGACCGCGCTCAGCGCTCAGAGCGCTACCTGGGCCATC	420
Db	361	TGCACCTACCGCAGCTGTGCATGACCGCGCTCAGCGCTCAGAGCGCTACCTGGGCCATC	420
Qy	421	TGCCGCCCGCTTCGGCGCCGCGCTTGGTCAACCGGGCGCGCGTTCGCGCGGCTCATCGCT	480
Db	421	TGCCGCCCGCTTCGGCGCCGCGCTTGGTCAACCGGGCGCGCGTTCGCGCGGCTCATCGCT	480
Qy	481	GTGCTCTGGGCGGTGGGGCTGTCTCTGTGCGGCTCCCTCTTGTTCGTGGGGGCTCGAG	540
Db	481	GTGCTCTGGGCGGTGGGGCTGTCTCTGTGCGGCTCCCTCTTGTTCGTGGGGGCTCGAG	540

Qy	541	CAGACCCCGGATCTCCGTAGTCCCGGGCTCAATGGCAACGGCGGATCGCTCTCTCG	600
Db	541	CAGACCCCGGATCTCCGTAGTCCCGGGCTCAATGGCAACGGCGGATCGCTCTCTCG	600
Qy	601	CTCTCGCCCTCGTCCGCGCTCTCTGCTCTCGGGGCGCACCGCGCTCCCGCGCTCG	660
Db	601	CTCTCGCCCTCGTCCGCGCTCTCTGCTCTCGGGGCGCACCGCGCTCCCGCGCTCG	660
Qy	661	GGGCGCGAGACCGCGAGGCGCGCGCTGTTACGCGCGGAATGCCCGCGGAGCCCGCG	720
Db	661	GGGCGCGAGACCGCGAGGCGCGCGCTGTTACGCGCGGAATGCCCGCGGAGCCCGCG	720
Qy	721	CAGTGGGCGGCTGCTGTATCTGTGGGTACCAACCGCTACTTCTTCTGCGCCCTT	780
Db	721	CAGTGGGCGGCTGCTGTATCTGTGGGTACCAACCGCTACTTCTTCTGCGCCCTT	780
Qy	781	CTGTGCTCAGCATCTCTACGGCTCATCGGGGAGTGTGGAGACGCGCGCGCG	840
Db	781	CTGTGCTCAGCATCTCTACGGCTCATCGGGGAGTGTGGAGACGCGCGCGCGCG	840
Qy	841	CTCGAGGCGCGCGCTCTCGGGGCGGAGAGGCGCACCGGAGACCGCTCGCTCTG	900
Db	841	CTCGAGGCGCGCGCTCTCGGGGCGGAGAGGCGCACCGGAGACCGCTCGCTCTG	900
Qy	901	CGTAAGTGGAGCCCGCTGTTTCAAGAAGCCCTGCTGCAGTCCGCGCGCGGGAGC	960
Db	901	C-----	901
Qy	961	CGCAAGCGTGGTCCCTTCCCTGCTGCCAGCTCTGGGCGCGCTTCCAGTCCC	1020
Db	902	-----	901
Qy	1021	TTTCTATTTCGATTCAGCTCCACCGCGTGGTGGTCTTGGCAITTAATTTG	1080
Db	902	-----TGGTGGTGGTCTTGGCAITTAATTTG	929
Qy	1081	CTGTTCCCTTCCAGTTCGAGAACTTATCATAAACAGGAAGATTGCGGATGAT	1140
Db	930	CTGTTCCCTTCCAGTTCGAGAACTTATCATAAACAGGAAGATTGCGGATGAT	989
Qy	1141	GTACTTCTCAGTACCTTACATCGTCTGCTGCACTTTCTATCTGAGCGCATCTAT	1200
Db	990	GTACTTCTCAGTACCTTAAACATCGTCTGCTGCACTTTCTATCTGAGCGCATCTAT	1049
Qy	1201	CAACCCATCTCTACAACTCTATTCAAGAAGTACAGAGCGCGCTTTAACTGCT	1260
Db	1050	CAACCCATCTCTACAACTCTATTCAAGAAGTACAGAGCGCGCTTTAACTGCT	1109
Qy	1261	GCTCGAAGGAAGTCCAGGCGGAGAGGCTTCCAGAAAGAGGACACTCGGGGGAAGT	1320
Db	1110	GCTCGAAGGAAGTCCAGGCGGAGAGGCTTCCAGAAAGAGGACACTCGGGGGAAGT	1169
Qy	1321	TGAGGGGACACTGGAGAGACAGGTTGGGTACACCGAGACAGCGCTAACGTGAAGAC	1380
Db	1170	TGAGGGGACACTGGAGAGACAGGTTGGGTACACCGAGACAGCGCTAACGTGAAGAC	1229
Qy	1381	GATGGGATAA 1390	
Db	1230	GATGGGATAA 1239	

RESULT 5
AF034632
LOCUS AF034632 2040 bp DNA linear PRI 13-JUL-1998
DEFINITION Homo sapiens orphan G protein-coupled receptor (GPR38) gene, complete cds.
ACCESSION AF034632
VERSION AF034632.1 GI:2654158
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE	1	(bases 1 to 2040)	
AUTHORS	McKee, K.K., Tan, C.P., Palyha, O.C., Liu, J., Feighner, S.D., Hreniuk, D.L., Smith, R.G., Howard, A.D. and Van Der Ploeg, L.H.		
TITLE	Cloning and characterization of two human G protein-coupled receptor genes (GPR38 and GPR39) related to the growth hormone secretagogue and neurotensin receptors		
JOURNAL	Genomics 46 (3), 426-434 (1997)		
MEDLINE	98110578		
PUBMED	9441746		
REFERENCE	2	(bases 1 to 2040)	
AUTHORS	McKee, K.K., Tan, C.P., Palyha, O.C., Liu, J., Feighner, S.D., Hreniuk, D.L., Smith, R.G., Van Der Ploeg, L.H.T. and Howard, A.D.		
TITLE	Direct Submission		
JOURNAL	Submitted (17-NOV-1997) Biochemistry and Physiology, Merck and Co., Inc., PO Box 2000, Rahway, NJ 07065, USA		
FEATURES	Location/Qualifiers		
source	1..2040		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
	/chromosome="13"		
	/map="13q14-21"		
	<1..>2040		
gene	/gene="GPR38"		
mRNA	join(<1..901,1703..>2040)		
	/gene="GPR38"		
CDS	join(1..901,1703..2040)		
	/gene="GPR38"		
	/note="orphan G protein-coupled receptor related to growth hormone secretagogue and neurotensin receptor"		
	/codon_start=1		
	/product="GPR38"		
	/protein_id="AAC26081.1"		
	/db_xref="GI:2654159"		
	/translation="MGSPNGSDGPEGAREPPWALPPCDERCSPPPLGALVPTAV		
	CICLFVVGSGNVTVMILGIRYDMRTTNTLYGSMVADLLILGLPFDLYRLMRSR		
	PWFPGFLCLRLSYGEGCTYATLLHMTALSVERYLAICRPLRLVLTNRVRLALIA		
	VLMVAALLAGPFLFLVGVQDQPGISVVFGLNGTARIASSPLASSPPLMLSRAPPSP		
	PSGPETAEEAALFSPRECRPAQGLALRVMLVTTAYFPFLFCLSLVGLIGLELWS		
	SRPLGPAASGRERHROTVRLLVVLAFITICWLPFHVGRIIYINTEDSRMYFSQ		
	YFNIVALQIPLYLSASINPLIYNLISKYRAAPFLKLLARKSRPGRHRSRDTAGEVAG		
	DTGGTGTGTGTTSTANVTMG"		
BASE COUNT	347 a 631 c 598 g 464 t		
ORIGIN			
Query Match	75.7%; Score 1052; DB 9; Length 2040;		
Best Local Similarity	100.0%; Pred. No. 1.1e-130; Mismatches 0; Indels 0; Gaps 0;		
Matches 1052; Conservative	0; Mismatches 0; Indels 0; Gaps 0;		
Qy	1	ATGGGAGAGCCCTTGGAAACGCGACGCGCGCGGCGCGCGCGCGCGCGCGCGCGCG	60
Db	1	ATGGGAGAGCCCTTGGAAACGCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	60
Qy	61	GCCTCGCGCTTTCGACGAGCGCGCTGCTGCTGCTTTCCTTTCCTTTCCTTTCCT	120
Db	61	GCCTCGCGCTTTCGACGAGCGCGCTGCTGCTGCTTTCCTTTCCTTTCCTTTCCT	120
Qy	121	GTGACCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	180
Db	121	GTGACCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	180
Qy	181	ATGCTGATCGGGCGCTACCGGACATGCGGACACCAACCACTTGTACCTTGGGCGCATG	240
Db	181	ATGCTGATCGGGCGCTACCGGACATGCGGACACCAACCACTTGTACCTTGGGCGCATG	240
Qy	241	GCGGTGTCGACCTACTCATCTGCTGCGGCTCGCGTTCGACCTGACCGCTCTGCGCG	300
Db	241	GCGGTGTCGACCTACTCATCTGCTGCGGCTCGCGTTCGACCTGACCGCTCTGCGCG	300
Qy	301	TCCGCGCCCTTGGGTGTTTCCGGCGCTGCTGCTGCGCGCTGCTGCTGCTGCTGCTG	360
Db	301	TCCGCGCCCTTGGGTGTTTCCGGCGCTGCTGCTGCGCGCTGCTGCTGCTGCTGCTG	360

Qy 361 TGCACCTACGCCACGCTGTGCACATGACCGCGCTCAGCGTCGAGCGCTACCTGGCCATC 420
 Db 361 TGCACCTACGCCACGCTGTGCACATGACCGCGCTCAGCGTCGAGCGCTACCTGGCCATC 420
 Qy 421 TGCACCGCGCTCCCGCGCGCGCTTGTGTACCGCGCGCGCGCTCCCGCGCGCTCATCGCT 480
 Db 421 TGCACCGCGCTCCCGCGCGCGCTTGTGTACCGCGCGCGCGCTCCCGCGCGCTCATCGCT 480
 Qy 481 GTGCTCTGGGCGCGTGGCGCTCTCTCTCGCGCGTCCCTTCTTGTCTGTGGTGGCGTCCGAG 540
 Db 481 GTGCTCTGGGCGCGTGGCGCTCTCTCTCGCGCGTCCCTTCTTGTCTGTGGTGGCGTCCGAG 540
 Qy 541 CAGGACCCCGGCATCTCCGTAGTCCCGGGCTCAATGGCACCGCGCGGATCGCTCCCTCG 600
 Db 541 CAGGACCCCGGCATCTCCGTAGTCCCGGGCTCAATGGCACCGCGCGGATCGCTCCCTCG 600
 Qy 601 CCTCTCGCTGTCTCGCGCTCTCTCTGTCTCTCGCGCGCGCGCGCGCGCGCGCGCTCG 660
 Db 601 CCTCTCGCTGTCTCGCGCTCTCTCTGTCTCTCGCGCGCGCGCGCGCGCGCGCGCTCG 660
 Qy 661 GGGCCCGAGACCGCGGAGCGCGCGCTGTTCAGCGCGCAATGCGCGCGCGCGCGCGCG 720
 Db 661 GGGCCCGAGACCGCGGAGCGCGCGCTGTTCAGCGCGCAATGCGCGCGCGCGCGCGCG 720
 Qy 721 CAGCTGGCGCGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 780
 Db 721 CAGCTGGCGCGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 780
 Qy 781 CTGTGCTCAGCATCTCTACGGCTCATCGCGCGCTCATCGCGCGGAGCTGTGAGCAGCGCGCGCG 840
 Db 781 CTGTGCTCAGCATCTCTACGGCTCATCGCGCGGAGCTGTGAGCAGCGCGCGCGCGCG 840
 Qy 841 CTGCGAGCGCGCGCGCTCGCGCGGAGAGAGGCGCACCGCGAGACCGTCCGCGTCTG 900
 Db 841 CTGCGAGCGCGCGCGCTCGCGCGGAGAGAGGCGCACCGCGAGACCGTCCGCGTCTG 900
 Qy 901 CGTAAGTGGAGCGCGCGCTGTTCAGAGAGCGCTGTCTGTCTGTCTGTCTGTCTGTCTGT 960
 Db 901 CGTAAGTGGAGCGCGCGCTGTTCAGAGAGCGCTGTCTGTCTGTCTGTCTGTCTGTCTGT 960
 Qy 961 GCGCAACCGTGGTGTCT 1020
 Db 961 GCGCAACCGTGGTGTCT 1020
 Qy 1021 TTTCTATTTCGATTCAGCTCCACCGCGCG 1052
 Db 1021 TTTCTATTTCGATTCAGCTCCACCGCGCG 1052

RESULT 6
 AL137000
 LOCUS
 DEFINITION Human DNA sequence from clone RP11-203116 on chromosome 13 Contains the gene for KIAA0970 protein, COX7CP1 (cytochrome c oxidase subunit VIIC pseudogene 1), a novel pseudogene, the GPR38 (G protein-coupled receptor 38) gene, ESTs, STSS, GSSs and a CpG island, complete sequence.
 AL137000
 ACCESSION
 VERSION AL137000.6 GI:9944121
 KEYWORDS HTG; COX7CP1; CpG island; cytochrome c oxidase; G protein-coupled receptor; GPR38; KIAA0970.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 163284)
 Direct Submission
 Submitted (18-DEC-2000) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 AUTHORS requests: clones@clones.sanger.ac.uk
 JOURNAL On Aug 29, 2000 this sequence version replaced gi:9926419.
 COMMENT During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 13, constructed by the Sanger Centre Chromosome 13 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr13>
 IMPORTANT: This sequence is not the entire insert of clone RP11-203116. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.
 The true left end of clone RP11-452110 is at 1 in this sequence. The true right end of clone RP11-103118 is at 163284 in this sequence. This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated repeat sequence elements. Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key. RP11-203116 is from the library RP11-11.1 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>
 VECTOR: pBACe3.6.

FEATURES

source Location/Qualifiers
 1..163284
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="13"
 /clone="RP11-203116"
 /clone_lib="RPCI-11.1"
 3..121
 /note="MER20 repeat: matches 4..126 of consensus"
 repeat_region
 152..261
 /note="MIR repeat: matches 20..137 of consensus"
 repeat_region
 262..554
 /note="AluX repeat: matches 1..293 of consensus"
 repeat_region
 555..698
 /note="MIR repeat: matches 137..255 of consensus"
 repeat_region
 1012..1291
 /note="AluJo repeat: matches 7..291 of consensus"
 repeat_region
 1320..1425
 /note="MER94 repeat: matches 4..114 of consensus"
 repeat_region
 1591..1804
 /note="L1PA8 repeat: matches 5945..6158 of consensus"
 repeat_region
 3240..3521
 /note="L2 repeat: matches 2310..2691 of consensus"
 repeat_region
 3824..4106
 /note="AluJo repeat: matches 1..301 of consensus"
 repeat_region
 4975..5042
 /note="L2 repeat: matches 2639..2705 of consensus"
 repeat_region
 6113..6408
 /note="AluX repeat: matches 1..295 of consensus"
 repeat_region
 7646..7732
 /note="AluSc repeat: matches 1..85 of consensus"
 repeat_region
 9887..10191
 /note="AluX repeat: matches 1..306 of consensus"
 repeat_region
 12201..12247
 /note="L2 repeat: matches 2704..2750 of consensus"
 repeat_region
 12472..12738
 /note="L2 repeat: matches 1304..1530 of consensus"
 repeat_region
 12739..12874
 /note="FLAM_C repeat: matches 2..133 of consensus"
 repeat_region
 12875..12887
 /note="L2 repeat: matches 1530..1542 of consensus"
 repeat_region
 12922..13227
 /note="L1MB3 repeat: matches 5871..6183 of consensus"
 repeat_region
 13303..13348

repeat_region	/note="MUT1E repeat: matches 1. .46 of consensus" 13533. .13623
repeat_region	/note="MUT1E repeat: matches 223. .315 of consensus" 13657. .13865
repeat_region	/note="MUT1E repeat: matches 307. .533 of consensus" 13890. .14150
repeat_region	/note="MUT1J repeat: matches 195. .456 of consensus" 14151. .14450
repeat_region	/note="AUsG repeat: matches 1. .307 of consensus" 14501. .14729
repeat_region	/note="L1M4C repeat: matches 588. .813 of consensus" 14730. .15006
repeat_region	/note="A1u0 repeat: matches 1. .285 of consensus" 15007. .15361
repeat_region	/note="L1M4C repeat: matches 813. .1178 of consensus" 15389. .15437
repeat_region	/note="L1M4C repeat: matches 1159. .1207 of consensus" 15438. .16412
repeat_region	/note="HERVL repeat: matches 2891. .3879 of consensus" 16413. .16616
repeat_region	/note="MUT2B repeat: matches 1. .216 of consensus" 17092. .17141
repeat_region	/note="MER5B repeat: matches 27. .79 of consensus" 17416. .17647
repeat_region	/note="M1R repeat: matches 17. .261 of consensus" 19402. .19499
repeat_region	/note="HY3 repeat: matches 2. .99 of consensus" 20730. .21163
repeat_region	/note="L2 repeat: matches 2168. .2609 of consensus" 25800. .25825
repeat_region	/note="13 copies 2 mer tt 92% conserved" 26041. .26131
repeat_region	/note="FLAM C repeat: matches 17. .107 of consensus" 26144. .26249
repeat_region	/note="53 copies 2 mer tt 64% conserved" 27029. .27259
repeat_region	/note="M1R repeat: matches 2. .251 of consensus" 27422. .27788
repeat_region	/note="L2 repeat: matches 1879. .2259 of consensus" 27863. .28239
repeat_region	/note="L2 repeat: matches 1278. .1734 of consensus" 28255. .28571
repeat_region	/note="A1uY repeat: matches 1. .310 of consensus" 28729. .29128
repeat_region	/note="L1M3 repeat: matches 4905. .5326 of consensus" 29170. .29413
repeat_region	/note="L1M3 repeat: matches 5461. .5715 of consensus" 29448. .29681
repeat_region	/note="L2 repeat: matches 1594. .1828 of consensus" 29683. .29886
repeat_region	/note="L1ME repeat: matches 5527. .5734 of consensus" 29888. .30323
repeat_region	/note="L2 repeat: matches 2129. .2596 of consensus" 30324. .30620
repeat_region	/note="A1uJB repeat: matches 1. .299 of consensus" 30621. .30780
repeat_region	/note="L2 repeat: matches 2596. .2749 of consensus" 30784. .30813
repeat_region	/note="15 copies 2 mer aa 86% conserved" 30814. .30947
repeat_region	/note="L1M3 repeat: matches 5724. .5874 of consensus" 31314. .33214
repeat_region	/note="M1R repeat: matches 175. .262 of consensus" 35454. .35753
misc_feature	/note="A1uS repeat: matches 1. .300 of consensus" 36245. .36743
repeat_region	/note="match: GSS: Em:AQ593229" 36856. .36843
misc_feature	/note="M1R repeat: matches 14. .237 of consensus" 38014. .38191
repeat_region	/note="match: STS: Em:G04228" 38279. .38416
repeat_region	/note="MER5A repeat: matches 7. .152 of consensus" 38279. .38416

repeat_region	39603..39708	/note="MIR repeat: matches 34. .144 of consensus"
repeat_region	40037..40103	/note="L2 repeat: matches 2664. .2731 of consensus"
repeat_region	40370..40523	/note="L2 repeat: matches 2562. .2701 of consensus"
repeat_region	40324..40810	/note="Alusx repeat: matches 2. .297 of consensus"
repeat_region	40811..41080	/note="L2 repeat: matches 2242. .2562 of consensus"
repeat_region	41085..41156	/note="L2 repeat: matches 2626. .2697 of consensus"
repeat_region	41349..41871	/note="MER2 repeat: matches 1. .343 of consensus"
misc_feature	complement(42041..42676)	/note="match: GSS: Em:AQ112495"
repeat_region	42192..48502	/note="L1PA7 repeat: matches 128. .6143 of consensus"
repeat_region	49671..49794	/note="L2 repeat: matches 2617. .2748 of consensus"
repeat_region	50285..50483	/note="L1MB3 repeat: matches 5985. .6183 of consensus"
repeat_region	50630..50815	/note="MIR repeat: matches 10. .203 of consensus"
repeat_region	50979..51656	/note="L2 repeat: matches 855. .1521 of consensus"
repeat_region	51654..51892	/note="L2 repeat: matches 1702. .1927 of consensus"
repeat_region	51893..52189	/note="Alusg repeat: matches 1. .297 of consensus"
repeat_region	52190..52379	/note="L2 repeat: matches 1927. .2153 of consensus"
Query Match 75.7%; Score 1052; DB 9; Length 163284;		
Best Local Similarity 100.0%; Pred.No.3.7e-131;		
Matches 1052; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	1	ATGGGCAGCCCTCGAAGCGGACGAGCGCCCGGAGGGGCGCGGAGCGCGCTGGCCCC 60
Db	144348	ATGGGCAGCCCTCGAAGCGGACGAGCGCCCGGAGGGGCGCGGAGCGCGCTGGCCCC 144407
Qy	61	GCGCTGCGCCTTCGCGAGCGCGCTGCTCGCCCTTTCCTCGGGGCGCTGTCGCG 120
Db	144408	GCGCTGCGCCTTCGCGAGCGCGCTGCTCGCCCTTTCCTCGGGGCGCTGTCGCG 144467
Qy	121	GTGACCGCTGTGTCCTGTGCTGTTCTGTCGTGGGGTGAGCGGCAACGTGTGACCGTGTG 180
Db	144468	GTGACCGCTGTGTCCTGTGCTGTTCTGTCGTGGGGTGAGCGGCAACGTGTGACCGTGTG 144527
Qy	181	ATGCTGATCGGGCGCTACCGGACATCGGACACCAACCACTTGTACTGGGACGACATG 240
Db	144528	ATGCTGATCGGGCGCTACCGGACATCGGACACCAACCACTTGTACTGGGACGACATG 144587
Qy	241	GCCGTGTCGACCTACTCATCTGCTCGGGGTGCGGTTTCGACCTCTACCGCTCTTCGGCGC 300
Db	144588	GCCGTGTCGACCTACTCATCTGCTCGGGGTGCGGTTTCGACCTCTACCGCTCTTCGGCGC 144647
Qy	301	TCGCGGCCCTGGGTGTTTCGGGCCGCTGCTCTGCGCCCTGTCCCTCTACGTGGCGAGGGC 360
Db	144648	TCGCGGCCCTGGGTGTTTCGGGCCGCTGCTCTGCGCCCTGTCCCTCTACGTGGCGAGGGC 144707
Qy	361	TGCACCTACGCCACGCTGCTGCATGACCGGCTCAGGCTCAGGCTACTCTGGCCATC 420
Db	144708	TGCACCTACGCCACGCTGCTGCATGACCGGCTCAGGCTCAGGCTACTCTGGCCATC 144767
Qy	421	TGCCGCCGCTTCGCGGCCGCGGTCTTGCTGTCACCGGCGCGGCTTCGGCGGCTCATCGCT 480
Db	144768	TGCCGCCGCTTCGCGGCCGCGGTCTTGCTGTCACCGGCGCGGCTTCGGCGGCTCATCGCT 144827
Qy	481	GTGCTCTGGGCGGTGGGCGTGTCTCTGCGCGGTTCCTCTTGTTCCTGGTGGCGCTCGAG 540
Db	144828	GTGCTCTGGGCGGTGGGCGTGTCTCTGCGGCTTCCTCTTGTTCCTGGTGGCGCTCGAG 144887

Db 320264 CAGACCCGCGCATCTCCGAGTCTCCGCGGCTCAATGGCACCGCGCGGATCGCTCTCTCG 320323
Qy 601 CCTCTCGCTCGTCCGCGCTCTCTGCTCTCGCGGCGGCACCGCGTCCCGCGCTCG 660
Db 320324 CCTCTCGCTCGTCCGCGCTCTCTGCTCTCGCGGCGGCACCGCGTCCCGCGCTCG 320383
Qy 661 GGGCCGAGACCGCGGAGCGCGCGCTGTTACGCGCGAATCCCGCGGAGCGCGCG 720
Db 320384 GGGCCGAGACCGCGGAGCGCGCGCTGTTACGCGCGAATCCCGCGGAGCGCGCG 320443
Qy 721 CAGTGGGCGGCTGCGCTGTCATGCTGTGGGTACACCGCTACTTCTTCTGCGCTTT 780
Db 320444 CAGTGGGCGGCTGCGCTGTCATGCTGTGGGTACACCGCTACTTCTTCTGCGCTTT 320503
Qy 781 CTGTGCTCAGCATCTCTACGGCTCATCGGGCTCATCGGGCGGAGCTGTGGAGCAGCGCGCG 840
Db 320504 CTGTGCTCAGCATCTCTACGGCTCATCGGGCGGAGCTGTGGAGCAGCGCGCG 320563
Qy 841 CTGCGAGCGCGCGCTCGGGCGGAGAGCGGCACCGGACAGCGTCCGCGTCTG 900
Db 320564 CTGCGAGCGCGCGCTCGGGCGGAGAGCGGCACCGGACAGCGTCCGCGTCTG 320623
Qy 901 CGTAAGTGGAGCGCGCTGTTCAAGAGCGCTGCTGAGTCCGCGCGCGCGGAGCC 960
Db 320624 CGTAAGTGGAGCGCGCTGTTCAAGAGCGCTGCTGAGTCCGCGCGCGCGGAGCC 320683
Qy 961 GCGCAACCGTGGTCCCTTCCCTGCTGCCCGAGCTGTGGGCGCGCTTCCAGCTCCC 1020
Db 320684 GCGCAACCGTGGTCCCTTCCCTGCTGCCCGAGCTGTGGGCGCGCTTCCAGCTCCC 320743
Qy 1021 TTTCCTATTTCGATTCCAGCTCCACCGCG 1052
Db 320744 TTTCCTATTTCGATTCCAGCTCCACCGCG 320775

RESULT 8
AX711879
LOCUS AX711879 349980 bp DNA linear PAT 08-MAY-2003
DEFINITION Sequence 1 from Patent WO03000727.
ACCESSION AX711879
VERSION AX711879.1 GI:29787684
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Zhang, Y., Moffatt, M., Cookson, W. and Tinsley, J.O.
TITLE Atopy
JOURNAL Patent: WO 03000727-A 1 03-JAN-2003;
ISIS INNOVATION LIMITED (GB)
FEATURES
source
1. 349980
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="Original length of seq 1 to big: 410,846-Splitted
into-seq 1: from 000.001 to 349.980-seq 327: from 300.001
to 410.846-Original length of seq 71 to big:
379.652-Splitted into-seq 71: from 000.001 to 349.980-seq
328: from 300.001 to 379.652"
BASE COUNT 106097 a 69968 c 68717 g 105198 t
ORIGIN

Query Match 75.7%; Score 1052; DB 6; Length 349980;
Best Local Similarity 100.0%; Pred. No. 3.1e-131;
Matches 1052; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGGCGACCCCTCGAAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
Db 38016 ATGGGCGACCCCTCGAAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 38075

Qy 61 GCCTCGCGCTTGTGCAAGAGCGCGCTGCTGCTCCCTTTCCTCGGGGCGCTGTGTCGG 120
Db 38076 GCCTCGCGCTTGTGCAAGAGCGCGCTGCTGCTCCCTTTCCTCGGGGCGCTGTGTCGG 38135
Qy 121 GTACACGCTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 180
Db 38136 GTACACGCTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 38195
Qy 181 ATGCTGATCGGGCGCTTACCGGAGATGCGGAGACCAACCAACCACTTGTACCTGGGAGCATG 240
Db 38196 ATGCTGATCGGGCGCTTACCGGAGATGCGGAGACCAACCAACCACTTGTACCTGGGAGCATG 38255
Qy 241 GCGGTGTGCAACCTACTACTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 300
Db 38256 GCGGTGTGCAACCTACTACTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 38315
Qy 301 TCGCGGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 360
Db 38316 TCGCGGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 38375
Qy 361 TGCACCTTAGCCCAACGCTGTGCTGACATGACCGCGCTCAGCGTCCGAGCGCTACCTGGCCATC 420
Db 38376 TGCACCTTAGCCCAACGCTGTGCTGACATGACCGCGCTCAGCGTCCGAGCGCTACCTGGCCATC 38435
Qy 421 TGCGGCGCGCTCGCGCGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 480
Db 38436 TGCGGCGCGCTCGCGCGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 38495
Qy 481 GTGCTGTGGGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 540
Db 38496 GTGCTGTGGGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 38555
Qy 541 CAGGACCCGCGCATCTCCGAGTCCCGGCGCTCAATGGACACCGCGGATCGCTCTCTCG 600
Db 38556 CAGGACCCGCGCATCTCCGAGTCCCGGCGCTCAATGGACACCGCGGATCGCTCTCTCG 38615
Qy 601 CTTCTCGCTGTGCGCGCTCTCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 660
Db 38616 CTTCTCGCTGTGCGCGCTCTCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 38675
Qy 661 GGGCCGAGACCGCGGAGCGCGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 720
Db 38676 GGGCCGAGACCGCGGAGCGCGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 38735
Qy 721 CAGCTGTGGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 780
Db 38736 CAGCTGTGGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 38795
Qy 781 CTGTGCTCAGCATCTCTTACCGGCTCATGCGGCGGAGCTGTGGAGCAGCGCGCGCGCG 840
Db 38796 CTGTGCTCAGCATCTCTTACCGGCTCATGCGGCGGAGCTGTGGAGCAGCGCGCGCGCG 38855
Qy 841 CTGCGAGCGCGCGCGCTCGGGCGGAGAGAGCGGCACCGGAGACCTGTCCGGGCTCTG 900
Db 38856 CTGCGAGCGCGCGCGCTCGGGCGGAGAGAGAGCGGCACCGGAGACCTGTCCGGGCTCTG 38915
Qy 901 CGTAAGTGGAGCGCGCTGTGCTTCAAGAGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 960
Db 38916 CGTAAGTGGAGCGCGCTGTGCTTCAAGAGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 38975
Qy 961 GCGCAACCGTGGGTCCTTCCCTTCCCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1020
Db 38976 GCGCAACCGTGGGTCCTTCCCTTCCCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 39035
Qy 1021 TTTCCTATTTCGATTCCAGCTCCACCGCG 1052
Db 39036 TTTCCTATTTCGATTCCAGCTCCACCGCG 39067

RESULT 9
AX739961
LOCUS AX739961 349980 bp DNA linear PAT 08-MAY-2003
DEFINITION Sequence 1 from Patent WO03000296.


```
Db 80 GCGTGGCGGCTTGGACGAGCGCGTGTCTGCGCCCTTTCCCTTGGGGGCGCTGTGTCGG 139
Qy 121 GTGACCGCTGTGTGCTGTGCTGTCTGTCTGCTGGGGGTGAGCGCAACAGTGTGACCGTG 180
Db 140 GTGACCGCTGTGTGCTGTGCTGTCTGTCTGCTGGGGGTGAGCGCAACAGTGTGACCGTG 199
Qy 181 ATGCTGATCGGGCGCTACCGGGAATGCGGACACCAACAACTTGTACCTGGGAGCATG 240
Db 200 ATGCTGATCGGGCGCTACCGGGAATGCGGACACCAACAACTTGTACCTGGGAGCATG 259
Qy 241 GCGGTGCGGACCTACTCATCTGCTGCGGGTGGCTGCGCTGCGCTGCGCTGCGCTGCGGC 300
Db 260 GCGGTGCGGACCTACTCATCTGCTGCGGGTGGCTGCGCTGCGCTGCGCTGCGCTGCGGC 319
Qy 301 TCGGGCGGCTTGGGTGTTGGGGCGCTGTCTGCGCGCTGTCTGCGCTGCGCTGCGCTGCGGC 360
Db 320 TCGGGCGGCTTGGGTGTTGGGGCGCTGTCTGCGCGCTGTCTGCGCTGCGCTGCGCTGCGGC 379
Qy 361 TGCACTACGACCGCTGCTGACATGACCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGGC 420
Db 380 TGCACTACGACCGCTGCTGACATGACCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGGC 439
Qy 421 TCGCGCGGCTTGGGTGTTGGGGCGCTGTCTGCGCGCTGTCTGCGCTGCGCTGCGCTGCGGC 480
Db 440 TCGCGCGGCTTGGGTGTTGGGGCGCTGTCTGCGCGCTGTCTGCGCTGCGCTGCGCTGCGGC 499
Qy 481 GTGCTGTGGGGCGTGGGGTGTCTGCTGCGCGCTGTCTGCGCTGCGCTGCGCTGCGCTGCGGC 540
Db 500 GTGCTGTGGGGCGTGGGGTGTCTGCTGCGCGCTGTCTGCGCTGCGCTGCGCTGCGCTGCGGC 559
Qy 541 CAGGACCGCGCATCTCCGTAGTCCCGGCGCTCAATGGCACCGCGCGCATGCGCTCCCTCG 600
Db 560 CAGGACCGCGCATCTCCGTAGTCCCGGCGCTCAATGGCACCGCGCGCATGCGCTCCCTCG 619
Qy 601 CCTCTCGGCTGCTGCGCGCTGTCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 660
Db 620 CTTTGTGCTGCTGCGCGCTGTCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 674
Qy 661 GGGCGCGAGACCGCG 675
Db 675 GGGCGCGAGACCGCG 689
```

```
RESULT 11
HSA342408 692 bp DNA linear PRI 18-JUL-2002
LOCUS Homo sapiens genomic sequence surrounding Not1 site, clone
DEFINITION NRI-PB24C.
VERSION AJ342408
KEYWORDS AJ342408.1 GI:15886883
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 692)
AUTHORS Kutsenko, A.S., Gizatullin, R.Z., Al-Amin, A.N., Wang, F., Kvasha, S.M.,
Podowski, R.M., Matushkin, Y.G., Gyanchandani, A., Muravenko, O.V.,
Levitky, V.G., Kolchanov, N.A., Protopopov, A.I., Kashuba, V.I.,
Kiselev, L.L., Wasserman, W., Wahlstedt, C. and Zabarovsky, E.R.
NOTI flanking sequences: a tool for gene discovery and verification
of the human genome
Nucleic Acids Res. 30 (14), 3163-3170 (2002)
JOURNAL 22131767
MEDLINE 12136098
PUBMED 12136098
REFERENCE 2 (bases 1 to 692)
AUTHORS Zabarovsky, E.R.
TITLE Direct Submission
JOURNAL Submitted (16-MAY-2001) Microbiology and Tumorbiology Centre,
Karolinska Institute, Theorells vag, 3, Box 280, Stockholm 171 77,
Sweden
FEATURES Location/Qualifiers
```

```
source 1. .692
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="NRI-PB24C"
BASE COUNT 92 a 239 c 219 g 140 t 2 others
ORIGIN
Query Match 41.2%; Score 572.2; DB 9; Length 692;
Best Local Similarity 91.5%; Pred. No. 7.9e-67;
Matches 616; Conservative 0; Mismatches 55; Indels 2; Gaps 1;
Qy 1 ATGGGAGAGCCCTTGGAAACGCGACGCGCCCGCGAGGGGCGGAGCGCCGCTGGCC 60
Db 20 ATGGGAGAGCCCTTGGAAACGCGACGCGCCCGCGAGGGGCGGAGCGCCGCTGGCC 79
Qy 61 GCGCTGCGGCTTGGCAAGAGCGCGCTGTCTGCGCCCTTTCCCTTGGGGGCGCTGGTCCG 120
Db 80 GCGCTGCGGCTTGGCAAGAGCGCGCTGTCTGCGCCCTTTCCCTTGGGGGCGCTGGTCCG 139
Qy 121 GTGACCGCTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 180
Db 140 GTGACCGCTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 199
Qy 181 ATGCTGATCGGGCGCTACCGGGAATGCGGACACCAACAACTTGTACCTGGGAGCATG 240
Db 200 ATGCTGATCGGGCGCTACCGGGAATGCGGACACCAACAACTTGTACCTGGGAGCATG 259
Qy 241 GCGGTGCGGACCTACTCATCTGCTGCGGGTGGCTGCGCTGCGCTGCGCTGCGCTGCGGC 300
Db 260 GCGGTGCGGACCTACTCATCTGCTGCGGGTGGCTGCGCTGCGCTGCGCTGCGCTGCGGC 319
Qy 301 TCGGGCGGCTTGGGTGTTGGGGCGCTGTCTGCGCGCTGTCTGCGCTGCGCTGCGCTGCGGC 360
Db 320 TCGGGCGGCTTGGGTGTTGGGGCGCTGTCTGCGCGCTGTCTGCGCTGCGCTGCGCTGCGGC 379
Qy 361 TGCACTACGACCGCTGCTGACATGACCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGGC 420
Db 380 TGCACTACGACCGCTGCTGACATGACCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGGC 439
Qy 421 TCGCGCGGCTTGGGTGTTGGGGCGCTGTCTGCGCGCTGTCTGCGCTGCGCTGCGCTGCGGC 480
Db 440 TCGCGCGGCTTGGGTGTTGGGGCGCTGTCTGCGCGCTGTCTGCGCTGCGCTGCGCTGCGGC 499
Qy 481 GTGCTGTGGGGCGTGGGGTGTCTGCTGCGCGCTGTCTGCGCTGCGCTGCGCTGCGCTGCGGC 540
Db 500 GTGCTGTGGGGCGTGGGGTGTCTGCTGCGCGCTGTCTGCGCTGCGCTGCGCTGCGCTGCGGC 559
Qy 541 CAGGACCGCGCATCTCCGTAGTCCCGGCGCTCAATGGCACCGCGCGCATGCGCTCCCTCG 600
Db 560 CATGACACCTGTATCTCCGTAGTCTCGGGCTCAATGCCACCGCGCGCATGCGCTTCTTC 619
Qy 601 CCT--CTCGGCTGCTGCGCGCTGTCTGCTGCGGGGCGCGCGCGCGCGCGCGCGCGCT 658
Db 620 GCTTATTAGCATAGTCGACGACTCTGTGTATATCGGGGTGCAACCGAGCATACATACCGT 679
Qy 659 CGGGGCGCGGAGAC 671
Db 680 CAGGGCGCGGATAC 692
RESULT 12
HSA339459
LOCUS Homo sapiens genomic sequence surrounding Not1 site, clone
DEFINITION NRI-WC14C.
VERSION AJ339459
KEYWORDS AJ339459.1 GI:15883877
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 692)
AUTHORS Kutsenko, A.S., Gizatullin, R.Z., Al-Amin, A.N., Wang, F., Kvasha, S.M.,
Podowski, R.M., Matushkin, Y.G., Gyanchandani, A., Muravenko, O.V.,
Levitky, V.G., Kolchanov, N.A., Protopopov, A.I., Kashuba, V.I.,
Kiselev, L.L., Wasserman, W., Wahlstedt, C. and Zabarovsky, E.R.
NOTI flanking sequences: a tool for gene discovery and verification
of the human genome
Nucleic Acids Res. 30 (14), 3163-3170 (2002)
JOURNAL 22131767
MEDLINE 12136098
PUBMED 12136098
REFERENCE 2 (bases 1 to 692)
AUTHORS Zabarovsky, E.R.
TITLE Direct Submission
JOURNAL Submitted (16-MAY-2001) Microbiology and Tumorbiology Centre,
Karolinska Institute, Theorells vag, 3, Box 280, Stockholm 171 77,
Sweden
FEATURES Location/Qualifiers
```

REFERENCE	1	(bases 1 to 615)			
AUTHORS	Kutsenko,A.S., Gizatullin,R.Z., Al-Amin,A.N., Wang,F., Kvasha,S.M., Podowski,R.M., Matushkin,Y.G., Gyanchandani,A., Muravenko,O.V., Levitsky,V.G., Kolchanov,N.A.I., Protopenov,A.I., Kashuba,V.I., Kiselev,L.L., Wasserman,W., Wahlestedt,C. and Zabarovsky,E.R.				
TITLE	Net1 flanking sequences: a tool for gene discovery and verification of the human genome				
JOURNAL	Nucleic Acids Res. 30 (14), 3163-3170 (2002)				
MEDLINE	22131767				
PUBMED	12136098				
REFERENCE	2	(bases 1 to 615)			
AUTHORS	Zabarovsky,E.R.				
TITLE	Direct Submission				
JOURNAL	Submitted (16-MAY-2001) Microbiology and Tumorbiology Centre, Karolinska Institute, Theorells vag, 3, Box 280, Stockholm 171 77, Sweden				
FEATURES	Location/Qualifiers				
source	1..615				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
	/clone="NR1-WC14C"				
BASE COUNT	61 a 224 c 204 g 121 t 5 others				
ORIGIN					
Query Match	40.1%; Score 557.4; DB 9; Length 615;				
Best Local Similarity	95.6%; Pred.No. 7.6e-65;				
Matches	570; Conservative 0; Mismatches 26; Indels 0; Gaps 0;				
Qy	1 ATGGGAGCCCTCGGAACGGCAGCGACGCCCGAGGGGGCGGGAGCGCCGTGGCCC 60				
Db	20 ATGGGAGCCCTCGGAACGGCAGCGACGCCCGAGGGGGCGGGAGCGCCGTGGCCC 79				
Qy	61 GCGCTGCCGCTTCGCAGCAGCGCGCTGCTCGCCCTTCCCTTGSGGGGCGTGTGCCG 120				
Db	80 GCGCTGCCGCTTCGCAGCAGCGCGCTGCTCGCCCTTCCCTTGSGGGGCGTGTGCCG 139				
Qy	121 GTACCCTGTGTGCTGTGCTCTTCCTGCTCGGGGTGAGCGGCAACGTGGTGACCGTG 180				
Db	140 GTACCCTGTGTGCTGTGCTGTTCGTGCTCGGGGTGAGCGGCAACGTGGTGACCGTG 199				
Qy	181 ATGCTGATCGGGCGCTACCGGGACATCGGACACCAACAATTGTACTTGGGACGATG 240				
Db	200 ATGCTGATCGGGCGCTACCGGGACATCGGACACCAACAATTGTACTTGGGACGATG 259				
Qy	241 GCCTGTCCGACCTACTCATCTGCTCGGGCTGCGCTTCGACCTGTACCGCTCTGGCG 300				
Db	260 GCCTGTCCGACCTACTCATCTGCTCGGGCTGCGCTTCGACCTGTACCGCTCTGGCG 319				
Qy	301 TCGCGGCCCTTGGGTGTTCGGGCCCTGCTCTCGCCCTGTCCCTCTACGTGGGCGAGGGC 360				
Db	320 TCGCGGCCCTTGGGTGTTCGGGCCCTGCTCTCGCCCTGTCCCTCTACGTGGGCGAGGGC 379				
Qy	361 TGCACTACGCCACCGCTGTGCTGCAATACACCGCGCTCAGCTCGAGCGCTACCTGGCCATC 420				
Db	380 TGCACTACGCCACCGCTGTGCTGCAATACACCGCGCTCAGCTCGAGCGCTACCTGGCCATN 439				
Qy	421 TGC CGCGCTTCCGCGCCCGCGCTTTGGTTCACCGGGCGCGCTCCGCGCGCTCATCGCT 480				
Db	440 TGNCGCGCTTCCGCGCCCGCGTTTGGGTTCACCGGGCGCGCTTNCGCGCGCTCATCGCT 499				
Qy	481 GTGCTCTGGGCGGTGGCGCTGCTCTCGCGGCTCCCTTTCTTGTTCGTGGGCGCTCGAG 540				
Db	500 GTGCTCTGGGCGGTGGCGCTGTGCTTTCGCGTCCCCCTTTTGTTCGTGGGCGGTNA 559				
Qy	541 CAGAACCCCGGCATCTCCGTAGTCCCGGGCTCAAATGGCACCGCGCGGATCGGCTC 596				
Db	560 CAAGACCCCGCATTTCCGTAGTNC CGGCTTAATGSCACCGCGGATATGCTC 615				
RESULT 13	HSA326768	701 bp	DNA	linear	PRI 18-JUL-2002
LOCUS	HSA326768				

DEFINITION: Homo sapiens genomic sequence surrounding NotI site, clone NR5-IP16C.	
ACCESSION	AJ326768
VERSION	AJ326768.1
KEYWORDS	GI:15971186
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1 (bases 1 to 701)
AUTHORS	Kutsenko, A.S., Gizatullin, R.Z., Al-Amin, A.N., Wang, F., Kvassha, S.M., Podowski, R.M., Matushkin, Y.G., Gyanchandani, A., Muravenko, O.V., Levitsky, V.G., Kolchanov, N.A., Protodopov, A.I., Kashuba, V.I., Kisselev, L.I., Wasserman, W., Wahlested, C. and Zabarovsky, E.R.
TITLE	NotI flanking sequences: a tool for gene discovery and verification of the human genome
JOURNAL	Nucleic Acids Res. 30 (14), 3163-3170 (2002)
MEDLINE	22131767
PUBMED	12136098
REFERENCE	2 (bases 1 to 701)
AUTHORS	Zabarovsky, E.R.
TITLE	Direct Submission
JOURNAL	Submitted (16-MAY-2001) Microbiology and Tumorbiology Centre, Karolinska Institute, Theorells vag, 3, Box 280, Stockholm 171 77, Sweden
FEATURES	Location/Qualifiers
source	1..701
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/clone="NR5-IP16C"
BASE COUNT	108 a 218 c 235 g 139 t 1 others
ORIGIN	
Query Match 30.6%; Score 424.8; DB 9; Length 701;	
Best Local Similarity 81.1%; Pred. No. 3e-47;	
Matches 554; Conservative 0; Mismatches 123; Indels 6; Gaps 5;	
Qy	1 ATGGGAGCCCTTGGAAACGCGACGACGGCCCGAGGGGGGCGGAGCGCGCGTGGCGCC 60
Db	20 ATGGGAGCCCTTGGAAACGCGACGCGCGCCCGAGGGGGGCGGAGCGCGCGGAGGCC 79
Qy	61 GCCTCTCGCGCTTGGACAGCGCCGCTGCTGCCCTTTCCCTTGGGGGGCGTGGTGC 120
Db	80 GCCTACCGACATGCTACGATCGCCGTTGCTCTCTCTTCCCTTGGGGGACATGGTGC 139
Qy	121 GTGACCGCTGTGTGCTGTGCTGTTCGTCTGTCGCGGGTGAGCGGCAACGTGTGACG 180
Db	140 GTGACCGCTGTGTGCTGTGCTCTTTCGTCTGTCGCGGGTGAGCAGCAACGTGTGAC 199
Qy	181 ATGCTGATCGGGCGCTACCGGGACATGCGGACACACAC -CAACTCTTACTCTGGG 239
Db	200 ATGATAACCGNCGCTCCCGGACATGCGGACACACACAGACTCTTACCTGTGCA 259
Qy	240 GGCGGTGTTCGACCTACTCATCTCTGCTCGGGGTGCGGTTTCGACCTGTACCGCT 299
Db	260 GGACAGGTTCACAACTACTTATCTGTCTCGGGTTCGCTTAGACCTAGACACGCT 319
Qy	300 CTCGGCGCCTGGGTGTTTCGGGCCGTGTCTGTGCGCGCTGTCCCTCTACGTGGG 359
Db	320 CTCGGGACCTGGGAGACGGGCCGTTGTCTGTGACGCTGTCCCTCTACTGTTGG 379
Qy	360 CTCGACCTAGCCACGCTGCTGCACATGACCGGCTCAGGCTCAGGCGTACCTGGCC 419
Db	380 CTCGACCTACTCCACGCTCTGGACATGACCGCGGTTCAGGCTAGAGCGTCTAC 439
Qy	420 CTCGCGCCGCTCCGCGCCGCTCTTGTGTTCACCCGGCGCGGTCGCGCGCTCATCG 479
Db	440 TTGCCACCGATCCGACCCGCGTTTGTGTACCCAGAGCCGAGTTTCGCGCGCTTAT 499
Qy	480 TGTGCTCTGGGCGTGGCGCTGCTCTCTCGCGTCCCTTCTTGTTCCTGGTGGCGTC 539
Db	500 TGTGCTCTGGGCGAGTGTCTGTCTATGGCGGTACCTTATGCTTGTGGGGTT - A 557

```

Qy 540 GCAGGACCCCGGATCTCCGTAGTCCGGCGCTCAATGGGACC-GCCGGGATCGCTCCT 598
Db 558 AGCAGACCCCGGATATGGTAGTACCGGGCTTAATGGCACCAGGCGGATCGTCTCCT 617
Qy 599 CGCTCTCGCTCGT-CGCCGCTCTCTGGCTCTCGGGGGCGCC-ACCGCGTCCCGGCC 656
Db 618 GGAATATAGTTTGTGGCCACATGTTGGTAGTGGCGGGCCCAACCGACGCTCCACAC 677
Qy 657 GTGGGGCCCGAGACCGCGGAGG 679
Db 678 GGTGTGCACACAGACCGCGGAGG 700

RESULT 14
AC109158/c
LOCUS
DEFINITION
AC109158
VERSION
AC109158.3 GI:31416079
KEYWORDS
HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 192116)
Birren,B., Nusbaum,C. and Lander,E.
Mus musculus, clone RP24-199E5
Unpublished
2 (bases 1 to 192116)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeArallano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C.,
Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T.,
Miengua,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (03-FEB-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 192116)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collymore,A., Cook,A., Cooke,P., Corum,B., DeArallano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Maclean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Meldrim,J., Meneus,L., Mihova,T., Miengua,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,

```

```

Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (05-JUN-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 5, 2003 this sequence version replaced gi:20303716.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence submissions@genome.wi.mit.edu
----- Project Information
Center project name: L18656
Center clone name: 199_E_5
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 188041 bases at least Q40
Consensus quality: 189224 bases at least Q30
Consensus quality: 189863 bases at least Q20
Insert size: 176000; agarose-fp
Insert size: 190616; sum-of-contigs
Quality coverage: 11.9 in Q20 bases; agarose-fp
Quality coverage: 11.0 in Q20 bases; sum-of-contigs
-----

```

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 16 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

```

```

1 291: contig of 291 bp in length
292 391: gap of 100 bp
392 1310: contig of 919 bp in length
1311 1410: gap of 100 bp
1411 2084: contig of 674 bp in length
2085 2184: gap of 100 bp
2185 3934: contig of 1750 bp in length
3935 4034: gap of 100 bp
4035 7860: contig of 3826 bp in length
7861 10973: contig of 3013 bp in length
10974 11073: gap of 100 bp
11074 13977: contig of 2904 bp in length
13978 13977: gap of 100 bp
14078 26266: contig of 12189 bp in length
26267 26366: gap of 100 bp
26367 36455: contig of 10089 bp in length
36456 36555: gap of 100 bp
36556 47992: contig of 11437 bp in length
47993 48092: gap of 100 bp
48093 66740: contig of 18647 bp in length
66740 66839: gap of 100 bp
66840 88335: contig of 21496 bp in length
88336 88435: gap of 100 bp
88436 124376: contig of 35941 bp in length
124377 124476: gap of 100 bp
124477 145657: contig of 21181 bp in length
145658 145757: gap of 100 bp
145758 165431: contig of 19674 bp in length
165432 165531: gap of 100 bp
165532 192116: contig of 26585 bp in length.
1.192116

```

FEATURES
source

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 31, 2003, 13:48:11 ; Search time 426.676 Seconds
(without alignments)
8794.080 Million cell updates/sec

Title: US-09-719-485-4
Perfect score: 1390
Sequence: 1 atggcagcccttggaacgg.....acgtgaagcagatgggataa 1390

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_19Jun03.*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
- 25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1390	100.0	1390	21	AAZ45404
2	1161	83.5	1161	22	AAF83684
3	1078	77.6	1239	21	AAZ45403
4	1078	77.6	1239	22	AAF83683
5	1078	77.6	1239	22	AAF85449
6	1078	77.6	1239	24	ABK90132
7	1078	77.6	1239	25	AB242842
8	1078	77.6	1506	22	AAI66989

9	1047.2	75.3	2040	21	AAA46116	Human G protein co
10	1038	74.7	3066	21	AAZ45402	Genomic sequence o
11	754.6	54.3	1203	22	AAF85448	Nucleotide sequenc
12	546	39.3	813	22	AAF85447	Nucleotide sequenc
13	501.2	36.1	1179	24	ABQ47146	Oligonucleotide fo
14	501.2	36.1	1179	24	ABQ47147	Oligonucleotide fo
15	437.2	31.5	1179	24	ABQ47148	Oligonucleotide fo
16	437.2	31.5	1179	24	ABQ47149	Oligonucleotide fo
17	283	20.4	283	19	AAV44930	Galanin receptor G
18	283	20.4	283	19	AAV32651	Galanin receptor G
19	283	20.4	283	19	AAV28290	Galanin receptor G
20	283	20.4	283	24	ABK14060	Rat galanin recept
21	248.2	17.9	1850	21	AAZ61492	cDNA encoding canl
22	234.8	16.9	1063	18	AAZ68662	Pig growth hormone
23	234.8	16.9	1063	18	AAZ69754	Swine growth hormo
24	233.2	16.8	1029	18	AAZ68663	Pig growth hormone
25	233.2	16.8	1029	18	AAZ69755	Swine growth hormo
26	231.6	16.7	1095	21	AAZ45993	cDNA encoding the
27	231.6	16.7	4009	21	AAZ45967	DNA encoding the m
28	229	16.5	250	25	ACA55761	Pig signalling pat
29	228.8	16.5	870	25	ABZ42674	Human growth hormo
30	228.8	16.5	1122	18	AAZ68665	Human growth hormo
31	228.4	16.4	1092	22	AAZ27800	Rat growth hormone
32	228.4	16.4	3129	18	AAZ68667	Rat growth hormone
33	228.4	16.4	3129	18	AAZ69759	Rat growth hormone
34	227.8	16.4	1088	18	AAZ68664	Human growth hormo
35	227.8	16.4	1088	18	AAZ69756	Human growth hormo
36	227.8	16.4	1101	21	AAA30643	Human G protein-co
37	227.8	16.4	1101	21	AAA30732	DNA encoding human
38	227.8	16.4	1101	21	AAZ51463	Human G protein-co
39	227.8	16.4	1101	22	AAF83680	Human G-protein co
40	227.8	16.4	1101	24	AAZ30395	Human growth-hormo
41	227.8	16.4	1101	25	ACA56550	Human signalling p
42	227.2	16.3	1122	18	AAZ69757	Human growth hormo
43	226.8	16.3	1092	18	AAZ69760	Rat growth hormone
44	225.2	16.2	1892	21	AAZ45405	DNA encoding the p
45	225.2	16.2	1092	22	AAF85450	Nucleotide sequenc

ALIGNMENTS

RESULT 1
AAZ45404
ID AAZ45404 standard; cDNA; 1390 BP.

XX AAZ45404;

AC AAZ45404;

XX 27-MAR-2000 (first entry)

XX cDNA encoding the motilin receptor splice variant MTL-R1B.

XX Motilin receptor; G-protein coupled receptor; MTL-R1; GPR38;
XX spliced form; MTL-R1a; MTL-R1b; gastric motility disorder;
XX functional defect; neurological disorder; scleroderma; colonoscopy;
XX paraneoplastic syndrome; radiation induced dysmotility; diabetes;
XX infection; stress-related motility disorder; psychogenic disorder;
XX gastroparesis; gastro-oesophageal reflux disease; constipation;
XX chronic idiopathic pseudo obstruction; acute faecal impaction;
XX postoperative ileus; gallstones; infantile colic; diarrhoea;
XX irritable bowel syndrome; non-ulcer dyspepsia; non-cardiac chest pain;
XX endoscopy; duodenal intubation; ds.

XX Homo sapiens.

XX Key Location/Qualifiers
XX CDS 1..1161
XX /*tag= a
XX /product= "MTL-R1B"

XX WO9964436-A1.

XX 16-DEC-1999.

PD

RESULT 3

```

AAZ45403
ID  AAZ45403 standard; cDNA; 1239 BP.
XX  AC  AAZ45403;
XX  DT  27-MAR-2000 (first entry)
XX  DE  cDNA encoding the motilin receptor splice variant MTL-R1A.
XX  KW  Motilin receptor; G-protein coupled receptor; MTL-R1; GPR38;
KW  spliced form; MTL-R1A; MTL-R1B; gastric motility disorder;
KW  functional defect; neurological disorder; scleroderma; colonoscopy;
KW  paraneoplastic syndrome; radiation induced dysmotility; diabetes;
KW  infection; stress-related motility disorder; psychogenic disorder;
KW  gastroparesis; gastro-oesophageal reflux disease; constipation;
KW  chronic idiopathic pseudo obstruction; acute faecal impaction;
KW  postoperative ileus; gallstones; infantile colic; diarrhoea;
KW  irritable bowel syndrome; non-ulcer dyspepsion; non-cardiac chest pain;
KW  endoscopy; duodenal intubation; ds.
XX  OS  Homo sapiens.
XX  FH  Key Location/Qualifiers
FT  CDS 1..1239
FT  /tag= a "MTL-R1A"
FT  /product= "MTL-R1A"
XX  PN  WO9964436-A1.
XX  PD  16-DEC-1999.
XX  PF  08-JUN-1999; 99WO-US12773.
XX  PR  12-JUN-1998; 98US-0089098.
XX  PA  (MERI ) MERCK & CO INC.
XX  PI  Feighner SD, Patchett AA, Tan C, McKee K, MacNeil D, Howard AD;
PI  Pong S, Smith RG;
XX  WPI: 2000-105868/09.
XX  P-PSDB; AAY54145.
XX  DR  Novel receptor protein for screening compounds used in treating
PT  irritable bowel syndrome, constipation and other gastric conditions
XX  Claim 4; Fig 2; 44pp; English.
XX  CC  The present sequence encodes splice variant MTL-R1A of the motilin
CC  receptor. The gene encodes a G-protein coupled receptor, and is
CC  designated MTL-R1 (also GPR38). Two spliced forms of GPR38 exist,
CC  MTL-R1A (see AAY54145) and MTL-R1B (see AAY54146). MTL-R1A is a
CC  functional seven transmembrane domain form, and MTL-R1B is a truncated
CC  five transmembrane domain. The MTL-R1 proteins are used to identify
CC  agonists and antagonists which can be used for treating gastric motility
CC  disorders, functional defects, disorders secondary to neurological
CC  disorders e.g. scleroderma, paraneoplastic syndromes radiation induced
CC  dysmotility, diabetes, infections, stress-related motility disorders,
CC  psychogenic disorders, gastroparesis, gastro-oesophageal reflux disease,
CC  constipation, chronic idiopathic pseudo obstruction, acute faecal
CC  impaction, postoperative ileus, gallstones, infantile colic, irritable
CC  bowel syndrome, non-ulcer dyspepsion, non-cardiac chest pain and
CC  diarrhoea. They can also be used in the preparation for colonoscopy,
CC  endoscopy and duodenal intubation.
XX  SQ  Sequence 1239 BP; 172 A; 431 C; 397 G; 239 T; 0 other;
Query Match 77.6%; Score 1078; DB 21; Length 1239;
Best Local Similarity 89.1%; Pred. No. 2e-186;
Matches 1239; Conservative 0; Mismatches 0; Indels 151; Gaps 1;

```

```

QY 1 ATGGGAGCCCTGGAAACGACGACGCGCCGAGGGGCGCGGAGCGCGCGTGGCCC 60
DB |||||
DB 1 ATGGGAGCCCTGGAAACGACGACGCGCCGAGGGGCGCGGAGCGCGCGTGGCCC 60
QY 61 GGGCTGGCGCTTGGGACGAGCGCGCTGCTGGCCCTTTCCCTTGGGGGGGCTGGTGGCG 120
DB |||||
DB 61 GGGCTGGCGCTTGGGACGAGCGCGCTGCTGGCCCTTTCCCTTGGGGGGGCTGGTGGCG 120
QY 121 GTGACCGCTGTGGCTGTGCTGCTGCTGCTGGGGGTGAGGGCAACCTGGTGACCGTG 180
DB |||||
DB 121 GTGACCGCTGTGGCTGTGCTGCTGCTGCTGGGGGTGAGGGCAACCTGGTGACCGTG 180
QY 181 ATGCTGATCGGCGCTTACCGGACATGCGGACACCAACCACTGTGTACCTGGGACGATG 240
DB |||||
DB 181 ATGCTGATCGGCGCTTACCGGACATGCGGACACCAACCACTGTGTACCTGGGACGATG 240
QY 241 GCGGTGTCGACCTACTCATCTGCTGCGGCTGCGCTTGCAGCTGTACCGCTTGGCGC 300
DB |||||
DB 241 GCGGTGTCGACCTACTCATCTGCTGCGGCTGCGCTTGCAGCTGTACCGCTTGGCGC 300
QY 301 TCGCGCGCTGGGTGTTGGGGCGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
DB |||||
DB 301 TCGCGCGCTGGGTGTTGGGGCGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 361 TGCACCTACGCCACGCTGCTGCACATGACCGCGCTCAGCGCTCAGCGCTACCTGGGCATC 420
DB |||||
DB 361 TGCACCTACGCCACGCTGCTGCACATGACCGCGCTCAGCGCTCAGCGCTACCTGGGCATC 420
QY 421 TGCAGCGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB |||||
DB 421 TGCAGCGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 481 GTGCTGTGGCGCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
DB |||||
DB 481 GTGCTGTGGCGCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 541 CAGGACCCCGCATCTCCGCTAGTCCCGGCTCAATGCGCACCGCGGATCGCTCCTCG 600
DB |||||
DB 541 CAGGACCCCGCATCTCCGCTAGTCCCGGCTCAATGCGCACCGCGGATCGCTCCTCG 600
QY 601 CTTCTGCTGCTGCGCGCTTCTGCTGCTTCCGCGGCGCCACCGCGCTCCCGCGCTCG 660
DB |||||
DB 601 CTTCTGCTGCTGCGCGCTTCTGCTGCTTCCGCGGCGCCACCGCGCTCCCGCGCTCG 660
QY 661 GGGCCCGAGACCGCGGAGCGCGGCTGTTTCAGCGCGGATGCGCGCGGAGCGCGCG 720
DB |||||
DB 661 GGGCCCGAGACCGCGGAGCGCGGCTGTTTCAGCGCGGATGCGCGCGGAGCGCGCG 720
QY 721 CAGCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT 780
DB |||||
DB 721 CAGCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT 780
QY 781 CTGTGCTCAGCATCTTCTACCGGCTCATCTGCGGCGGAGCTGTGGAGCAGCGCGCGCG 840
DB |||||
DB 781 CTGTGCTCAGCATCTTCTACCGGCTCATCTGCGGCGGAGCTGTGGAGCAGCGCGCGCG 840
QY 841 CTGCGAGCGCGCGCTTCCGCGGCGGAGAGGCGCACCGGAGACCGCTTCCGCTGCTG 900
DB |||||
DB 841 CTGCGAGCGCGCGCTTCCGCGGCGGAGAGGCGCACCGGAGACCGCTTCCGCTGCTG 900
QY 901 CGTAAGTGAGCGCGCTGTTTCCAAAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
DB |||||
DB 901 C-----
QY 961 GCGCAACGCTGGGTCCCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
DB |||||
DB 902 -----
QY 1021 TTTCTATTTCGATTCAGCTCCACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
DB |||||
DB 902 -----
QY 1081 CTGGTTGCTTTCAGCTGGGAGATCATTTTACATAAACGGAAGATTCGCGGATGAT 1140

```



```

Db 930 CTGTTCCCTTCCAGTTCGAGATCATTTATATAAACACGGAAGATTCGGGATGAT 989
Qy 1141 GTACTTCTCAGTACTTTAAACATCGTCTGCTGCAACTTTTCTATCTGAGCGCATCTAT 1200
Db 990 GTACTTCTCAGTACTTTAAACATCGTCTGCTGCAACTTTTCTATCTGAGCGCATCTAT 1049
Qy 1201 CAACCCAAATCTCTACAACTCATTTCAAGAAGTACAGAGCGCGGCTTTAAACTGCT 1260
Db 1050 CAACCCAAATCTCTACAACTCATTTCAAGAAGTACAGAGCGCGGCTTTAAACTGCT 1109
Qy 1261 GCTCGCAAGAGTCCAGGCGGAGAGGCTTCCACAGAGCAGGACACTCGGGGGAGT 1320
Db 1110 GCTCGCAAGAGTCCAGGCGGAGAGGCTTCCACAGAGCAGGACACTCGGGGGAGT 1169
Qy 1321 TGACGGGGACACTGGAGGAGACAGCGTGGGCTACACCGAGACAGCCCTAAACGTGAAGAC 1380
Db 1170 TGCAGGGGACACTGGAGGAGACAGCGTGGGCTACACCGAGACAGCCCTAAACGTGAAGAC 1229
Qy 1381 GATGGGATAA 1390
Db 1230 GATGGGATAA 1239

```

RESULT 4

```

AAF83683
ID AAF83683 standard; DNA; 1239 BP.
AC AAF83683;
XX
DT 23-JUL-2001 (first entry)
XX
DE Long form of motilin receptor, GPR-38A isoform encoding DNA.
XX
KW zsig33; signal transduction; hormone; enzyme; neural development;
KW gastric contractility; nutrient uptake; digestive; pancreatic; human;
KW insulin-like growth factor-I; growth hormone; bone; gastrointestinal;
KW glucose; osteopathic; anorectic; vulnery; immunomodulator; GHS-R; ds;
KW G-protein coupled receptor; motilin receptor; GPR38; GPR-38A; isoform.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1239
FT FT /*tag= a
FT FT /product= "GPR-38A"
XX
PN W0200138355-A2.
XX
PD 31-MAY-2001.
XX
PF 22-NOV-2000; 2000WO-US32074.
XX
PR 22-NOV-1999; 99US-0166765.
XX
PA (ZYMO ) ZYMOGENETICS INC.
XX
PI Sheppard PO, Jaspers SR, Deisher TA, Bishop PD;
XX
DR WPI; 2001-355879/37.
DR P-PSDB; AAB62652.
XX
XX Forming reversible peptide receptor complex for purifying cell and
XX peptides, stimulating signal transduction and modulating hormone
XX secretion, involves contacting a receptor with zsig33 polypeptide
XX
PS Disclosure; Page 102-104; 11ipp; English.
XX
XX The invention relates to a method of forming a reversible peptide-
XX receptor complex that involves providing an immobilized peptide,
XX contacting the receptor with a zsig33 peptide (comprising residues 24-37
XX of AAB62649), where the receptor binds to the zsig33 peptide. The method
XX is useful for purifying cells, purifying a peptide, stimulating signal

```

```

CC transduction in a cell expressing a receptor. It is also useful for
CC modulating secretion of hormones, neural development and/or utilization,
CC gastric contractility, nutrient uptake, secretion of digestive and
CC pancreatic enzymes and hormones, secretion of insulin-like growth factor
CC -I, secretion of non-zsig33 proteins. It is useful for modulating growth
CC hormone secretion in a mammal having a disease associated with abnormal
CC levels of growth hormone, such as osteoporosis, bone repair, bone
CC remodeling, low osteoblast levels, cartilage repair and remodeling,
CC skeletal dysplasia, immune suppression, obesity, growth retardation,
CC protein catabolic responses after surgery, cachexia, protein loss,
CC dwarfism, wound healing and ovulation induction, treating a mammal having
CC a metabolic disorder requiring neurological feedback, such as satiety
CC regulation, glucose absorption and metabolism and neuropathy-associated
CC gastrointestinal disorders, and stimulating glucose-induced insulin
CC release in a mammal. The present sequence represents the DNA encoding
CC the long form of motilin receptor, GPR-38A (one of the two isoforms of
CC GPR38 which result from alternative splicing). GPR38 has homology to the
CC human G-protein coupled receptor, GHS-R.
XX
SQ Sequence 1239 BP; 172 A; 431 C; 397 G; 239 T; 0 other;

```

```

Query Match 77.6%; Score 1078; DB 22; Length 1239;
Best Local Similarity 89.1%; Pred. No. 2e-186;
Matches 1239; Conservative 0; Mismatches 0; Indels 151; Gaps 1;

```

```

Qy 1 ATGGGAGAGCCCTTGGAAACGACAGCAGCGCCCGAGGGGCGGGAGCGCCGTGGCCC 60
Db 1 ATGGGAGAGCCCTTGGAAACGACAGCAGCGCCCGAGGGGCGGGAGCGCCGTGGCCC 60
Qy 61 GCGCTGCGGCTTGGACAGAGCGCGCTGCTCGCCCTTTCCCTTGGGGGGCTGGTGGC 120
Db 61 GCGCTGCGGCTTGGACAGAGCGCGCTGCTCGCCCTTTCCCTTGGGGGGCTGGTGGC 120
Qy 121 GTACACCGCTGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 121 GTACACCGCTGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Qy 181 ATGCTGATCGGGCGCTACCGGAGCATGCGGACACCAACCACTTGTACCTGGGCGAG 240
Db 181 ATGCTGATCGGGCGCTACCGGAGCATGCGGACACCAACCACTTGTACCTGGGCGAG 240
Qy 241 GCGGTGTGCGACCTACTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Db 241 GCGGTGTGCGACCTACTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Qy 301 TCGGCGCCCTGGGTGTTGCGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Db 301 TCGGCGCCCTGGGTGTTGCGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Qy 361 TGCACCTTACGCCACCGCTGTGTCATGACCGCGCTCAGCGCTCAGCGGCTACCTGGCCATC 420
Db 361 TGCACCTTACGCCACCGCTGTGTCATGACCGCGCTCAGCGGCTCAGCGGCTACCTGGCCATC 420
Qy 421 TCGCGCGCGCTCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 421 TCGCGCGCGCTCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Qy 481 GTGCTCTGGGCGCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 481 GTGCTCTGGGCGCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Qy 541 CAGGACCCCGGCATCTCCGTAGTCCCGGGGCTCAATGGCACCGCGGATCGCCTCCCTCG 600
Db 541 CAGGACCCCGGCATCTCCGTAGTCCCGGGGCTCAATGGCACCGCGGATCGCCTCCCTCG 600
Qy 601 CCTCTCGCTCTGTCGCGCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 601 CCTCTCGCTCTGTCGCGCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Qy 661 GGGCCCGAGACCCCGGAGCGCGCGCTGCTTTCAGCGCGCAATGCGGCGGAGCCCGCG 720
Db 661 GGGCCCGAGACCCCGGAGCGCGCGCTGCTTTCAGCGCGCAATGCGGCGGAGCCCGCG 720

```

QY 721 CAGTGGGCGCGTGTGCTGTCATGCTGTGGGTACACCGCCTACTTCTTCTGCGCCTTT 780
 Db 721 CAGTGGGCGCGTGTGCTGTCATGCTGTGGGTACACCGCCTACTTCTTCTGCGCCTTT 780
 QY 781 CTGTGCTCAGCATCTCTACGGGCTCATCGGGGGAGCTGTGGAGCAGCCGGCGCG 840
 Db 781 CTGTGCTCAGCATCTCTACGGGCTCATCGGGGGAGCTGTGGAGCAGCCGGCGCG 840
 QY 841 CTGGAGGCGCGCGCTCGGGGGAGAGAGCCACCGGAGACCGTCCGCGTCTG 900
 Db 841 CTGGAGGCGCGCGCTCGGGGGAGAGAGCCACCGGAGACCGTCCGCGTCTG 900
 QY 901 CGTAAGTGGAGCGCGCTGTGTTCCAAAGACCGCTGCTGCAAGTCCGCGCGCGGAGC 960
 Db 901 C----- 901
 QY 961 GCGCAACGCTGGGTCCCTTCCCTGCTGCCAGCTCTGGGCGCGCTTCCAGTCCC 1020
 Db 902 ----- 901
 QY 1021 TTTCTATTTCGATTCCAGCTCCACCGCGGTGGTGGTCTGGCATTTATATTG 1080
 Db 902 -----TGGTGGTGGTCTGGCATTTATATTG 929
 QY 1081 CTGTTGCCCTTCCAGCTTGGCAGAAATCAATTACATAAACAAGAAATTCGGGATGAT 1140
 Db 930 CTGTTGCCCTTCCAGCTTGGCAGAAATCAATTACATAAACAAGAAATTCGGGATGAT 989
 QY 1141 GTACTTCTCAGTACTTTAAACATCGTGGCTGTGCAACTTTTCTATCTGAGCGCATCTAT 1200
 Db 990 GTACTTCTCAGTACTTTAAACATCGTGGCTGTGCAACTTTTCTATCTGAGCGCATCTAT 1049
 QY 1201 CAACCCAACTCTTACAACTCTTCAAGAGAGTACAGAGCGCGCGCTTTAACTGCT 1260
 Db 1050 CAACCCAACTCTTACAACTCTTCAAGAGAGTACAGAGCGCGCGCTTTAACTGCT 1109
 QY 1261 GCTCGAAGAGTCCAGGCGCGAGAGCTTCCACAGAGAGAGGACACTCGCGGGGAAGT 1320
 Db 1110 GCTCGAAGAGTCCAGGCGCGAGAGCTTCCACAGAGAGAGGACACTCGCGGGGAGT 1169
 QY 1321 TGAGGGGACACTGGAGGAGACAGCTGGGCTACCGAGACAAAGCGCTAACGTGAAGAC 1380
 Db 1170 TGAGGGGACACTGGAGGAGACAGCTGGGCTACCGAGACAAAGCGCTAACGTGAAGAC 1229
 QY 1381 GATGGGATAA 1390
 Db 1230 GATGGGATAA 1239

RESULT 5

AAF85449 ID AAF85449 standard; cDNA; 1239 BP.

AC AAF85449;

XX 23-JUL-2001 (first entry)

DE Nucleotide sequence of a human motilin receptor polypeptide.

XX Motilin receptor; gastrointestinal disease; gastric motility disorder;
 KW gastroparesis; irritable bowel syndrome; diarrhoea; ss.
 XX Homo sapiens.

XX Key Location/Qualifiers
 FH 1..1239
 FT CDS /*tag= a
 FT /note= "motilin receptor"
 FT

XX W0200132710-A1..

XX 10-MAY-2001.

XX

PF 25-OCT-2000; 2000WO-US29426.
 XX 29-OCT-1999; 99US-0162264.
 XX (MERI) MERCK & CO INC.
 XX Tan C, McKee K;
 XX WPI; 2001-343479/36.
 DR P-PSDB; AAB68478.
 XX Novel polypeptides related to dog and rabbit motilin receptor
 PT polypeptide, comprising unique regions from dog and motilin receptor
 PT amino acid sequence, useful for identifying compounds for treating
 PT diarrhoea in humans
 XX Disclosure; Page 34; 42pp; English.
 XX The present sequence encodes a human motilin receptor polypeptide.
 CC The specification describes a unique sequence present in exon 1 of
 CC the dog motilin receptor, which is not present in human or Sphaeroides
 CC nephelus 7587 motilin receptor sequences. The unique nucleic acid
 CC sequence is useful for measuring the ability of a compound to affect
 CC motilin receptor activity. Motilin receptor polynucleotides and
 CC polypeptides are used to identify therapeutic compounds which are
 CC useful for treating gastrointestinal diseases and disorders such as
 CC gastric motility disorders, gastroparesis, irritable bowel syndrome,
 CC and diarrhoea.
 XX SQ Sequence 1239 BP; 172 A; 431 C; 397 G; 239 T; 0 other;
 Query Match 77.6%; Score 1078; DB 22; Length 1239;
 Best Local Similarity 89.1%; Pred. No. 2e-186; Indels 151; Gaps 1;
 Matches 1239; Conservative 0; Mismatches 0;
 QY 1 ATGGGAGAGCCCTTGAACCGGACGACGCGCCGAGGGGCGCGGAGCGCGCGTGGCCC 60
 Db 1 ATGGGAGAGCCCTTGAACCGGACGACGCGCCGAGGGGCGCGGAGCGCGCGTGGCCC 60
 QY 61 GCGCTCGCGCTTGGACAGAGCGCGCTGCTGCGCCCTTTCCCTGGGGGGCGTGTGCGG 120
 Db 61 GCGCTCGCGCTTGGACAGAGCGCGCTGCTGCGCCCTTTCCCTGGGGGGCGTGTGCGG 120
 QY 121 GTGACCGCTGTGTCGCTGTCGCTGTCGTCGCGGGGTGAGCGGCAACGTTGACCGGTG 180
 Db 121 GTGACCGCTGTGTCGCTGTCGCTGTCGTCGCGGGGTGAGCGGCAACGTTGACCGGTG 180
 QY 181 ATGCTGATCGGGCGCTTACCGGGACATGCGGACACCAACCACTTGTACCTGGGAGCATG 240
 Db 181 ATGCTGATCGGGCGCTTACCGGGACATGCGGACACCAACCACTTGTACCTGGGAGCATG 240
 QY 241 GCGGTGTCCGACCTACTCATCTGCTGCGGGCTGCGGTGTCAGCTGTACCGCCCTCTGGCGC 300
 Db 241 GCGGTGTCCGACCTACTCATCTGCTGCGGGCTGCGGTGTCAGCTGTACCGCCCTCTGGCGC 300
 QY 301 TCGCGGCGCTGGGTGTTGCGGCGCTGCTGCGCGCTGTCCTCTACGTTGGCGGAGGGC 360
 Db 301 TCGCGGCGCTGGGTGTTGCGGCGCTGCTGCGCGCTGTCCTCTACGTTGGCGGAGGGC 360
 QY 361 TGCACCTAGCCACCGCTGTGTCATGACCGCGCTCAGCGCTCAGCGCTTACCTGGCCATC 420
 Db 361 TGCACCTAGCCACCGCTGTGTCATGACCGCGCTCAGCGCTTACCTGGCCATC 420
 QY 421 TGCAGCGCGCTGCGCGCGCGCTGTTGGTCAACCGGCGCGCGCTGCGCGCGCTCATCGCT 480
 Db 421 TGCAGCGCGCTGCGCGCGCGCTGTTGGTCAACCGGCGCGCGCTGCGCGCGCTCATCGCT 480
 QY 481 GTGCTCTGGGCGGTGGGCTGCTCTGCGGGTCCCTTCTTCTGTTGGGCGGTGAG 540
 Db 481 GTGCTCTGGGCGGTGGGCTGCTCTGCGGGTCCCTTCTTCTGTTGGGCGGTGAG 540
 QY 541 CAGGACCCCGGCATCTCCGCTAGTCCCGGGCTCAATGGGACCGCGCGGATCGCTCCCTCG 600

Db 541 CAGACCCCGCATCTCGTAGTCCCGGCTCAATGGCAGCCGCGGATCGCTCTCTCG 600
Qy 601 CCTCTCGCTCGTCCGCGCTCTCTGCTCTCGCGGCGCCACCGCGTCCCGCGCTCG 660
Db 601 CCTCTCGCTCGTCCGCGCTCTCTGCTCTCGCGGCGCCACCGCGTCCCGCGCTCG 660
Qy 661 GGGCCGAGACCGGAGCGCGCGCTCTGCTCTCGCGGCGCCACCGCGTCCCGCGCTCG 720
Db 661 GGGCCGAGACCGGAGCGCGCGCTCTGCTCTCGCGGCGCCACCGCGTCCCGCGCTCG 720
Qy 721 CAGTGGGCGCGCTCGCTGCTCATCTGCTGCTCACCACCGCTCTCTCTCTCTCTCT 780
Db 721 CAGTGGGCGCGCTCGCTGCTCATCTGCTGCTCACCACCGCTCTCTCTCTCTCTCT 780
Qy 781 CTGTGCTCAGCATCTCTCTAGCGCTCATCGGCGGAGCTGTGAGCAGCGCGCGCG 840
Db 781 CTGTGCTCAGCATCTCTCTAGCGCTCATCGGCGGAGCTGTGAGCAGCGCGCGCG 840
Qy 841 CTGGAGCGCGCGCTCTGGGCGGAGAGGCGCACCGGCGAGCGTCCGCTCTG 900
Db 841 CTGGAGCGCGCGCTCTGGGCGGAGAGGCGCACCGGCGAGCGTCCGCTCTG 900
Qy 901 CGTAAGTGGAGCGCGCTGGTTCACAGAGCGCTGCTGAGTCCGCGCGCGGAGC 960
Db 901 C----- 901
Qy 961 GCGCAACGTGGGTCCCTTCCCTCTGCTCGCCAGCTCTGGGCGCGCTTCCAGTCCC 1020
Db 902 ----- 901
Qy 1021 TTTCCTATTTCGATTCCAGCTCCACCGCGGTGGTGGTCTGGCATTTATAATTG 1080
Db 902 -----TGGTGGTGGTCTGGCATTTATAATTG 929
Qy 1081 CTGTTGCCCTTCCAGCTTCGAGATCAATTAACAACACGGAATTCGCGATGAT 1140
Db 930 CTGTTGCCCTTCCAGCTTCGAGATCAATTAACAACACGGAATTCGCGATGAT 989
Qy 1141 GTACTTCTCAGTACTTAAACATCTGCTCTGCAACTTTTCTATCTGAGCGCATCTAT 1200
Db 990 GTACTTCTCAGTACTTAAACATCTGCTCTGCAACTTTTCTATCTGAGCGCATCTAT 1049
Qy 1201 CAACCAATCTCTAACCTCATTTCAAGAGTACAGAGCGCGCTTTAACTGCT 1260
Db 1050 CAACCAATCTCTAACCTCATTTCAAGAGTACAGAGCGCGCTTTAACTGCT 1109
Qy 1261 GCTGCAAGAGTCCAGGCGGAGGCTTCCAGAGCAGGACACTCGGCGGAGT 1320
Db 1110 GCTGCAAGAGTCCAGGCGGAGGCTTCCAGAGCAGGACACTCGGCGGAGT 1169
Qy 1321 TGCAGGGGACACTGGAGGAGACGCTGGGCTACACCGAGACGCTTAACGTGAAGAC 1380
Db 1170 TGCAGGGGACACTGGAGGAGACGCTGGGCTACACCGAGACGCTTAACGTGAAGAC 1229
Qy 1381 GATGGGATAA 1390
Db 1230 GATGGGATAA 1239

RESULT 6
ABK90132
ID ABK90132 standard; DNA; 1239 BP.
XX
AC ABK90132;
XX
DT 21-OCT-2002 (first entry)
XX
DE DNA encoding human G protein-coupled receptor 38 (GPR38).
XX
KW Human; G protein-coupled receptor 38; receptor; GPR38; gene; ds;
KW Alzheimer's disease; Parkinson's disease; ulcerative colitis;
KW Crohn's disease; Hodgkin's disease; glioblastoma; breast carcinoma;
KW colon carcinoma; lung small cell carcinoma; lung adenocarcinoma;

KW pancreatic small cell carcinoma; pancreatic adenocarcinoma.
XX Homo sapiens.
XX Location/Qualifiers
PH 1..1239
FT /tag= a
FT /product= "Human G protein-coupled receptor 38 (GPR38)"
PN WO200257791-A2.
XX
XX 25-JUL-2002.
XX 29-NOV-2001; 2001WO-US45219.
XX 29-NOV-2000; 2000US-250251P.
PR 30-NOV-2000; 2000US-250452P.
XX (LIFE-) LIFESPAN BIOSCIENCES INC.
XX Brown JP, Burner GC, Roush CL, Kulander BG;
XX WPI: 2002-566812/60.
XX P-PSDB; ABG30936.
XX Assay for detecting Alzheimer's disease, Parkinson's disease,
XX ulcerative colitis, Crohn's disease, Hodgkin's disease, glioblastoma or
XX carcinoma, comprises using a binding partner for G protein coupled
XX receptor 38 -
XX Disclosure; Fig 1; 112pp; English.
XX The present invention relates to a new assay method that involves
XX contacting a binding partner specific for G protein coupled receptor
XX (GPR) 38 with specific cells. The method of the invention is useful for
XX the detection of an increased risk of Alzheimer's disease, Parkinson's
XX disease, ulcerative colitis, Crohn's disease, Hodgkin's disease,
XX glioblastoma, or carcinoma. GPR 38 is used to manufacture a medicament
XX for inhibiting, treating or preventing Alzheimer's disease, Parkinson's
XX disease, ulcerative colitis, Crohn's disease, Hodgkin's disease,
XX glioblastoma, breast carcinoma, colon carcinoma, lung small cell
XX carcinoma, lung adenocarcinoma, pancreatic small cell carcinoma, and
XX pancreatic adenocarcinoma. An agonist or antagonist to GPR 38 are used
XX to manufacture a medicament able to reduce the symptoms of these
XX diseases. Nucleic acids encoding GPR 38 can also be used to treat the
XX diseases. The present nucleic acid sequence encodes the human G protein-
XX coupled receptor 38 (GPR38) of the invention.
XX Sequence 1239 BP; 172 A; 431 C; 397 G; 239 T; 0 other;

Query Match 77.6%; Score 1078; DB 24; Length 1239;
Best Local Similarity 89.1%; Pred. No. 2e-186;
Matches 1239; Conservative 0; Mismatches 0; Indels 151; Gaps 1;
Qy 1 ATGGCGAGCCCTTGGAAACGGCAGCGCCCGAGGGGCGCGGAGCGCGTGGCCC 60
Db 1 ATGGCGAGCCCTTGGAAACGGCAGCGCCCGAGGGGCGCGGAGCGCGTGGCCC 60
Qy 61 GCGCTGCGCGCTTGGCAGCGCCCTGCTGCGCCCTTTCCCTGGGGGCGCTGGTCCG 120
Db 61 GCGCTGCGCGCTTGGCAGCGCCCTGCTGCGCCCTTTCCCTGGGGGCGCTGGTCCG 120
Qy 121 GTGACCGCTGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 121 GTGACCGCTGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Qy 181 ATGCTGATCGGGCGCTACCGGGACATGCGGACACCAACCACTTGACCTGGGAGCATG 240
Db 181 ATGCTGATCGGGCGCTACCGGGACATGCGGACACCAACCACTTGACCTGGGAGCATG 240
Qy 241 GCGGTGTCGACCTACTCATCTGCTGCGGCTCGGCTCGGCTGACCTGTACCGCTTGGCGC 300
Db 241 GCGGTGTCGACCTACTCATCTGCTGCGGCTCGGCTCGGCTGACCTGTACCGCTTGGCGC 300

CC used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode
 CC GPCR proteins given in ABP81675 to ABP82018, which are used in the
 CC exemplification of the present invention.

XX	Sequence	1239 BP; 172 A; 431 C; 397 G; 239 T; 0 other;	
XX	Query Match	77.6%; Score 1078; DB 25; Length 1239;	
XX	Best Local Similarity	89.1%; Pred. No. 2e-186;	
XX	Matches 1239; Conservative	0; Mismatches 0; Indels 151; Gaps 1;	
QY	1	ATGGGAGAGCCCTGGAAACGACGAGCGCGCCCGAGGGGCGCGGAGCGCGCGTGGCC 60	
DB	1	ATGGGAGAGCCCTGGAAACGACGAGCGCGCCCGAGGGGCGCGGAGCGCGCGTGGCC 60	
QY	61	GCCTGCGCGCTTGGGAGAGCGCGCTGCTGCGCCCTTTCCCTTGGGGGCGCTGGTGGCG 120	
DB	61	GCCTGCGCGCTTGGGAGAGCGCGCTGCTGCGCCCTTTCCCTTGGGGGCGCTGGTGGCG 120	
QY	121	GTGACCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180	
DB	121	GTGACCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180	
QY	181	ATGCTGATCGGGGCGCTACCGGGACATGCGGACCAACCACTTGTACCTGGGACGATG 240	
DB	181	ATGCTGATCGGGGCGCTACCGGGACATGCGGACCAACCACTTGTACCTGGGACGATG 240	
QY	241	GCCTGTGCGACCTACTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300	
DB	241	GCCTGTGCGACCTACTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300	
QY	301	TGCGGCGCTGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360	
DB	301	TGCGGCGCTGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360	
QY	361	TGCACCTACGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420	
DB	361	TGCACCTACGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420	
QY	421	TGCGGCGCTGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480	
DB	421	TGCGGCGCTGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480	
QY	481	GTGCTGTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540	
DB	481	GTGCTGTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540	
QY	541	CAGGACCGCGCATCTCCGTAGTCCCGGGCTCAATGGCACCGGGGATCGCTCTCG 600	
DB	541	CAGGACCGCGCATCTCCGTAGTCCCGGGCTCAATGGCACCGGGGATCGCTCTCG 600	
QY	601	CTCTGCGCTGTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660	
DB	601	CTCTGCGCTGTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660	
QY	661	GGGCGCGAGACCGCGAGGCGCGCGCTGTTACGCGCGGAATGCGCGCGAGCGCGCG 720	
DB	661	GGGCGCGAGACCGCGAGGCGCGCGCTGTTACGCGCGGAATGCGCGCGAGCGCGCG 720	
QY	721	CAGTGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780	
DB	721	CAGTGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780	
QY	781	CTGTGCTCAGCATCTCTACGGGCTCATCGGGGAGCTGTGGAGCAGCGCGGCGCG 840	
DB	781	CTGTGCTCAGCATCTCTACGGGCTCATCGGGGAGCTGTGGAGCAGCGCGGCGCG 840	
QY	841	CTGCGAGCGCGCGCTGTGGGCGGAGAGGCGCACCGGAGACCGTCCGCGTCTG 900	
DB	841	CTGCGAGCGCGCGCTGTGGGCGGAGAGGCGCACCGGAGACCGTCCGCGTCTG 900	
QY	901	CGTAAGTGGAGCGCGCTGTGTTCAAAGACGCTGCTGCTGCTGCTGCTGCTGCTGCT 960	
DB	901	C-----	

RESULT 8

AAI66989
 ID AAI66989 standard; DNA; 1506 BP.

XX AAI66989;

XX AC

DT 30-JAN-2002 (first entry)

XX Human GPR38 variant GPR38V polypeptide encoding DNA.

XX GPR38V; variant; antibacterial; cytostatic; analgesic; antispasmodic;
 anti-Parkinsonian; hypertensive; hypotensive; antidiabetic; osteopathic;
 anti-allergic; anti-migraine; neuroleptic; nootropic; anticonvulsant;
 anti-ulcer; antiemetic; cardiac; vaccine; ds.

OS Homo sapiens.

PH Key Location/Qualifiers

FT CDS 1..1506

FT /*tag= a

FT /product= "GPR38V"

XX WO200164836-A2.

XX 07-SEP-2001.

XX 28-FEB-2001; 2001WO-US06277.

XX 01-MAR-2000; 2000US-0516315.

XX (SMIK) SMITHKLINE BEECHAM CORP.

XX Elshourbagy N, Shabon U;

XX WPI; 2001-638956/73.

XX P-PSDB; AAG5822.

XX New human GPR38V polypeptide and polynucleotide, useful for treating
 e.g. bacterial, fungal, protozoal and viral infections, cancers or

PT allergies, as vaccines, and for identifying agonists and antagonists
 PT potentially useful in therapy
 XX Claim 2; Page 26; 32pp; English.
 PS This DNA encodes a human GPR38 variant (GPR38V) polypeptide. GPR38V can
 XX be expressed by standard recombinant methodology. The polynucleotides and
 CC polypeptides are used in the treatment of bacterial, fungal, protozoal
 CC and viral infections, infections caused by HIV-1 or HIV-2, pain, cancers,
 CC diabetes, obesity, anorexia, asthma, Parkinson's disease, acute heart
 CC failure, hypertension, urinary retentions, osteoporosis, allergies,
 CC ulcers, migraine, psychotic and neurological disorders, or dyskinesias.
 CC They are also useful for identifying agonists and antagonists that are
 CC potentially useful in therapy, as vaccines to induce immunological
 CC response in a mammal. The polypeptides may also be used as immunogens to
 CC produce antibodies immunospecific for the polypeptides, and to identify
 CC membrane bound or soluble receptors.
 XX
 SQ Sequence 1506 BP; 207 A; 534 C; 509 G; 256 T; 0 other;
 Query Match 77.6%; Score 1078; DB 22; Length 1506;
 Best Local Similarity 89.1%; Pred. No. 2e-186;
 Matches 1239; Conservative 0; Mismatches 0; Indels 151; Gaps 1;
 1 ATGGGCGAGCCCTGGAAACGCGAGCGACGCGCCCGAGGGGCGGAGCGCGCGTGGCCC 60
 268 ATGGGCGAGCCCTGGAAACGCGAGCGACGCGCCCGAGGGGCGGAGCGCGCGTGGCCC 327
 61 GCGCTGCGCGCTTGGAGAGAGCGCGCTGCTGCGCCCTTTCCCTGGGGGGCGTGGTGGCG 120
 328 GCGCTGCGCGCTTGGAGAGAGCGCGCTGCTGCGCCCTTTCCCTGGGGGGCGTGGTGGCG 387
 121 GTACCGCTGTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
 388 GTACCGCTGTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 447
 181 ATGCTGATCGGGGCGCTACCGGGGACATCGGGGACACACCACTTGTACCTGGGCGAGCATG 240
 448 ATGCTGATCGGGGCGCTACCGGGGACATCGGGGACACACCACTTGTACCTGGGCGAGCATG 507
 241 GCGGTGTCCAGCTATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
 508 GCGGTGTCCAGCTATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 567
 301 TCGGGCGCTGTGGGTGTTGGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
 568 TCGGGCGCTGTGGGTGTTGGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 627
 361 TGACCTTACCGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
 628 TGACCTTACCGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 687
 421 TGCGGCGCGCTGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 688 TGCGGCGCGCTGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 747
 481 GTGCTGTGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 748 GTGCTGTGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 807
 541 CAGGACCCCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
 808 CAGGACCCCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 867
 601 CTTCTGCGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
 868 CTTCTGCGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 927
 661 GGGCCCGAGACCGGGAGCGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
 928 GGGCCCGAGACCGGGAGCGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 987
 721 CAGCTGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780

Db 988 CAGCTGGGGCGCTGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1047
 Qy 781 CTGTGCTCAGCATCTCTACGGGCTCATCGGGGAGCTGCGGAGCTGCGGAGCTGCGGAGCTG 840
 Db 1048 CTGTGCTCAGCATCTCTACGGGCTCATCGGGGAGCTGCGGAGCTGCGGAGCTGCGGAGCTG 1107
 Qy 841 CTGCGAGGCGCGGCGCGCTGCGGGGAGAGAGGCGCACCGGAGAGCTGCGGAGCTGCGGAGCTG 900
 Db 1108 CTGCGAGGCGCGGCGCGCTGCGGGGAGAGAGGCGCACCGGAGAGCTGCGGAGCTGCGGAGCTG 1167
 Qy 901 CGTAAGTGGAGCGCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
 Db 1168 C----- 1168
 Qy 961 GCGCAACGCTGGGTCCCTCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
 Db 1169 ----- 1169
 Qy 1021 TTTCTATTTCGATTCCAGCTCCACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
 Db 1169 -----TGGTGGTGGTCTGCGCATTTATATTG 1196
 Qy 1081 CTGTTGCGCTTCCACGTTGGCAGAAATCATTTATATAAACAACGGAAGATTTCGCGGATGAT 1140
 Db 1197 CTGTTGCGCTTCCACGTTGGCAGAAATCATTTATATAAACAACGGAAGATTTCGCGGATGAT 1256
 Qy 1141 GTACTTCTCTCAGTACTTTAAACATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
 Db 1257 GTACTTCTCTCAGTACTTTAAACATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1316
 Qy 1201 CAACCAATCTCTACAACTCATTTCAAAGAGTACAGAGCGCGCGCTTTTAAACTGCT 1260
 Db 1317 CAACCAATCTCTACAACTCATTTCAAAGAGTACAGAGCGCGCGCTTTTAAACTGCT 1376
 Qy 1261 GCTCGCAAGAAAGTTCAGCGCGAGAGGCTTCCACAGAGCAGGACACCTGCGGGGGAAGT 1320
 Db 1377 GCTCGCAAGAAAGTTCAGCGCGAGAGGCTTCCACAGAGCAGGACACCTGCGGGGGAAGT 1436
 Qy 1321 TGCAGGGGACACTGGAGGAGACACGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
 Db 1437 TGCAGGGGACACTGGAGGAGACACGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1496
 Qy 1381 GATGGGATAA 1390
 Db 1497 GATGGGATAA 1506
 RESULT 9
 ID AAA46116
 XX AAA46116 standard; cDNA; 2040 BP.
 AC AAA46116;
 XX
 DT 22-AUG-2000 (first entry)
 XX
 DE Human G protein coupled receptor hGPR38 (V297K) cDNA SEQ ID NO:129.
 XX
 KW Human; G protein coupled receptor; GPCR; transmembrane receptor;
 KW identification; agonist; screening; therapeutic; pharmaceutical;
 KW mutant; ss.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN WO200022131-A2.
 XX
 PD 20-APR-2000.
 XX
 PF 13-OCT-1999; 99WO-US24065.
 XX
 PR 13-OCT-1998; 98US-0170496.
 PR 12-NOV-1998; 98US-0108029.
 PR

OS	Oryctolagus cuniculus.		
XX			
PH	Key	Location/Qualifiers	
FT	CDS	1..1203	
FT		/*tag= a	
FT		/note= "motilin receptor"	
PN	WO200132710-A1.		
PN			
XX			
PD	10-MAY-2001.		
XX			
XX	25-OCT-2000; 2000WO-US29426.		
XX			
XX	29-OCT-1999; 99US-0162264.		
PR			
PA	(MERI) MERCK & CO INC.		
XX			
XX	Tan C, McKee K;		
XX			
DR	WPI; 2001-343479/36.		
DR	P-PSDB; AAB68477.		
XX			
PT	Novel polypeptides related to dog and rabbit motilin receptor		
PT	polypeptide, comprising unique regions from dog and motilin receptor		
PT	amino acid sequence, useful for identifying compounds for treating		
PT	diarrhoea in humans		
XX			
XX	Claim 18; Page 18-19; 42pp; English.		
PS			
XX			
CC	The present sequence encodes a rabbit motilin receptor polypeptide.		
CC	The specification describes a unique sequence present in exon 1 of		
CC	the dog motilin receptor, which is not present in human or Spinaeroides		
CC	nephelus 75E7 motilin receptor sequences. The unique nucleic acid		
CC	sequence is useful for measuring the ability of a compound to affect		
CC	motilin receptor activity. Motilin receptor polynucleotides and		
CC	polypeptides are used to identify therapeutic compounds which are		
CC	useful for treating gastrointestinal diseases and disorders such as		
CC	gastric motility disorders, gastroparesis, irritable bowel syndrome,		
CC	and diarrhoea.		
XX			
SQ	Sequence 1203 BP; 154 A; 423 C; 403 G; 223 T; 0 other;		
	Query Match 54.3%; Score 754.6; DB 22; Length 1203;		
	Best Local Similarity 76.2%; Pred. No. 7.6e-128;		
	Matches 1058; Conservative 0; Mismatches 144; Indels 187; Gaps 3;		
QY	1 ATGGGACGCCCTTGGACGCGCAGCGCGCCCGGCGGGGGCGGGAGCGCGCGTGGCCC	60	
DB	1 ATGGGACGCCCTTGGAAACGCGAGCGAGCGCGCCCGGAGGACGCGGGGCGCGTGGGCC	60	
QY	61 GCGCTGCGCGCTTGGACGAGCGCGCGTGTGCGCCCTTTCCCTTGGGGGGCGCTGGTGGCG	120	
DB	61 GCGCTGCGCGCGTGGATGAGCGCGTGTGCGCCCTTCCCTTGGGACGCTGGTGGCT	120	
QY	121 GTGACCGCTGTGCGCTGTGCTTGGTGTGCTGCGGGGTGAGCGGCAACGTTGTAACCGTG	180	
DB	121 GTGACCGCGCTGTGCGCTTGGCGCTTTCGCGGTGAGCGCAACGTTGGTGAACCGTG	180	
QY	181 ATGCTGATCGGGCGCTACCGGACATCGGACCAACCACTGTGACTTGGGCGAGCATG	240	
DB	181 CTGCTGATCGGGCGCTACCGGACATCGGACCAACCACTGTGACTTGGGCGAGCATG	240	
QY	241 GCGGTGTCGACCTACTCATCTCTGCTGCGGCTGCGGCTGTGACCTGTACCGCCCTTGGCGC	300	
DB	241 GCGGTGTCGACCTGTCTCTCTGCTGCGGCTGCGGCTGTGACCTGTACCGCCCTTGGCGC	300	
QY	301 TCGGGCGCTGTGGGTGTTTGGGCGCGCTGTCTGCGCGCTGTCCCTCTACGTGGGCGAGGC	360	
DB	301 TCGAGGCGCTGTGGGTGTTTGGGCGAGCTGTCTGCGCGCTGTCTGCTGTACGTGGGCGAGGC	360	
QY	361 TGCACCTACCGCAGCTGTGTCATGACATGACCGCGCTGAGCGCTACCTGGCGCATC	420	
DB	361 TGCACCTACCGCTGTGCTGTGTCATGACCGCGCTGAGCGCTACCTGGCGCATC	420	

RESULT 12
AAF85447
ID AAF85447 standard; DNA; 813 BP.

[illegible]


```
Qy 301 TCGGGCCCTGGGTGTTGGGGCCCTGCTCTGCCGCGCTGCTCTAGTGGGGGAGGGC 360
Db      |||||
Qy 742 TCGGGGTTTGGGTGTTTCGGGTGCTGTTTGTCTGTTTGTCTGTTTACGTGGGGAGGGT 801
Db      |||||
Qy 361 TGCACCTACGCCACGCTGCTGCACATACACGCGCTCAGCGTCGAGCGCTACCTGGCCATC 420
Db      |||||
Qy 802 TGTATTTACGTTAGTTGTTGTATATGATCGCGTTAGCGTTCGAGCGTTATTGGTTATT 861
Db      |||||
Qy 421 TGCCGCCGCTCCGCGCCCGCGCTTGGTCACCCGGCGCGCGTCCGCGGCTCATCGCT 480
Db      |||||
Qy 862 TGTGTTGCTTCGTTTCGCGTTTCGCGTTTGGTTATTCCGCGTCGCGTTTCGCGGTTTATCGTT 921
Db      |||||
Qy 481 GTGCTCTGGGCGCGTGGCGCTGCTCTCTCCGCTCCCTTCTTGTCTTGGTGGGCTCGAG 540
Db      |||||
Qy 922 GTGTTTGGGTGCTGGGCTTGTCTTGTCTGGTCTTGTCTGGTCTTGTCTGGTCTGGT 981
Db      |||||
Qy 541 CAGGACCCGCGCATCTCCGTAGTCCCGGCTCAATGGCACCGCGGATCGCCTCCTCG 600
Db      |||||
Qy 982 TAGGATTCGGTATTTTCGTAGTTTCGGTTTAAATGGTATCGCGCGGATCGTTTTCG 1041
Db      |||||
Qy 601 CCTCTCGCTCGTCCGCGCTCTCTGGCTCTCGCGGGCGCACCGCGTCCCGCGCTCG 660
Db      |||||
Qy 1042 TTTTCGTTTCGTCGTCGTTTTCGTTTTCGTTTTCGCGGCGTTATCGTCTTTCGTCGTCG 1101
Db      |||||
Qy 661 GGGCCCGAGACCGCGGAGCGCGCGCTGTTTACGCGCGAATCCCGCGGAGCCCCGCG 720
Db      |||||
Qy 1102 GGGTTCGAGATCGCGAGTCCGCGGCTTGTATTAGTCGCGAATGTCGCTCGAGTTTCGCG 1161
Db      |||||
Qy 721 CAGCTGGGCGCGTGCCT 738
Db      |||||
Qy 1162 TAGTTGGGCGGTTGCGT 1179
Db      |||||
```

Search completed: December 31, 2003, 22:22:45
Job time : 438.676 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 31, 2003, 21:43:42 ; Search time 98.3416 Seconds
(without alignments)
6238.690 Million cell updates/sec

Title: US-09-719-485-4
Perfect score: 1390
Sequence: 1 atgggcagcccttggaacgg.....acgtgaagacgatgggataa 1390

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	283	20.4	283	3	US-08-993-088A-4
2	283	20.4	283	4	US-08-993-424B-4
3	283	20.4	283	4	US-09-603-680-4
4	234.8	16.9	1063	3	US-09-077-675A-1
5	234.8	16.9	1063	4	US-09-077-674-1
6	230	16.5	1029	3	US-09-077-675A-4
7	230	16.5	1029	4	US-09-077-674-4
8	229	16.5	250	4	US-09-016-434-359
9	228.8	16.5	1122	3	US-09-077-675A-9
10	228.8	16.5	1122	4	US-09-077-674-9
11	228.4	16.4	1092	3	US-09-077-675A-15
12	228.4	16.4	1092	4	US-09-077-674-15
13	228.4	16.4	3129	3	US-09-077-675A-14
14	228.4	16.4	3129	4	US-09-077-674-14
15	227.8	16.4	1088	3	US-09-077-675A-6
16	227.8	16.4	1088	4	US-09-077-674-6
17	227.8	16.4	1101	4	US-09-016-434-1148
18	227.8	16.4	1101	4	US-09-170-496D-87
19	227.8	16.4	1101	4	US-09-170-496D-209
20	153	11.0	836	3	US-09-077-675A-11
21	153	11.0	836	4	US-09-077-674-11
22	134.8	9.7	1248	4	US-09-545-944-1
23	127	9.1	1575	3	US-08-858-876A-1
24	127	9.1	1575	3	US-09-472-880-1
25	125.4	9.0	1342	3	US-08-832-399-1
26	125.4	9.0	1342	3	US-09-372-498-1
27	123	8.8	4080	4	US-09-016-434-1346

Sequence 3, Appl1
Sequence 3, Appl1
Sequence 1366, Ap
Sequence 12, Appl
Sequence 113, App
Sequence 223, App
Sequence 3, Appl1
Sequence 3, Appl1
Sequence 6, Appl1
Sequence 6, Appl1
Sequence 6, Appl1
Sequence 27, Appl1
Sequence 3, Appl1
Sequence 18, Appl
Sequence 19, Appl
Sequence 18, Appl
Sequence 18, Appl
Sequence 19, Appl

28 122.2 8.8 1529 3 US-08-858-876A-3
29 122.2 8.8 1529 3 US-09-472-880-3
30 122 8.8 1285 4 US-09-016-434-1366
31 120 8.6 1535 4 US-09-668-680-12
32 119 8.6 1212 4 US-09-170-496D-113
33 119 8.6 1212 4 US-09-170-496D-223
34 115.4 8.3 1776 4 US-09-200-090-3
35 115.4 8.3 1233 4 US-09-200-090-1
36 109.8 7.9 1164 3 US-08-993-088A-6
37 109.8 7.9 1164 4 US-08-993-424B-6
38 109.8 7.9 1164 4 US-09-603-680-6
39 109.8 7.9 1365 4 US-08-899-112B-27
40 108.2 7.8 1219 4 US-08-981-700A-3
41 106.8 7.7 1116 3 US-08-993-088A-18
42 106.8 7.7 1116 3 US-08-993-088A-19
43 106.8 7.7 1116 4 US-08-993-424B-18
44 106.8 7.7 1116 4 US-09-603-680-18
45 106.8 7.7 1116 4 US-09-603-680-19

ALIGNMENTS

RESULT 1
US-08-993-088A-4
; Sequence 4, Application US/08993088A
; Patent No. 6287855
; GENERAL INFORMATION:
; APPLICANT: Tan, Carina
; APPLICANT: Sullivan, Kathleen
; TITLE OF INVENTION: GALANIN RECEPTOR GALR2 AND
; NUCLEOTIDES ENCODING SAME
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/993,088A
; FILING DATE: 18-DEC-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/033,851
; FILING DATE: 27-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Heber, Sheldon O.
; REGISTRATION NUMBER: 38,179
; REFERENCE/DOCKET NUMBER: 19846
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1958
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 283 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Other
; FEATURE:
; NAME/KEY: Other
; LOCATION: 1...283
; OTHER INFORMATION: cdna probe
US-08-993-088A-4

Query Match 20.4%; Score 283; DB 3; Length 283;
 Best Local Similarity 100.0%; Pred. No. 2.6e-48;
 Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 206 TGGGACCAACCACTTGTACCTGGGAGCATGGCCGTGTCGACCTACTATCTGTC 265
 Db 1 TGGGACCAACCACTTGTACCTGGGAGCATGGCCGTGTCGACCTACTATCTGTC 60

QY 266 TCGGGCTGCGGTTGACCTGTACCGCTCTGGGGCTCGCGCCCTGGGGCGC 325
 Db 61 TCGGGCTGCGGTTGACCTGTACCGCTCTGGGGCTCGCGCCCTGGGGCGC 120

QY 326 TGCTCTCGCGCTGTCCCTCTACGTGGGCGAGGCTGCACCTAGCCACGCTGTGCACA 385
 Db 121 TGCTCTCGCGCTGTCCCTCTACGTGGGCGAGGCTGCACCTAGCCACGCTGTGCACA 180

QY 386 TGACCGGCTCAGGCTCGAGGCTACCTGGCCATCTGCCGCCGCTCGCGCGCGGTCT 445
 Db 181 TGACCGGCTCAGGCTCGAGGCTACCTGGCCATCTGCCGCCGCTCGCGCGCGGTCT 240

QY 446 TGGTACCCGCGCGCGCTCGCGCGCTCATCGCTGTGCTCTG 488
 Db 241 TGGTACCCGCGCGCGCTCGCGCGCTCATCGCTGTGCTCTG 283

RESULT 2
 US-09-93-424B-4
 ; Sequence 4, Application US/08993424B
 ; Patent No. 6337206
 ; GENERAL INFORMATION:
 ; APPLICANT: Tan, Carina
 ; APPLICANT: Kolakowski, Lee F., Jr.
 ; TITLE OF INVENTION: MOUSE GALANIN RECEPTOR GALR2 AND
 ; TITLE OF INVENTION: NUCLEOTIDES ENCODING SAME
 ; NUMBER OF SEQUENCES: 18
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Merck & Co., Inc.
 ; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
 ; CITY: Rahway
 ; STATE: NJ
 ; COUNTRY: USA
 ; ZIP: 07065-0900
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: Windows
 ; SOFTWARE: FASTSEQ For Windows Version 2.0b
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/993,424B
 ; FILING DATE: 18-DEC-1997
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/033,851
 ; FILING DATE: 27-DEC-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Heber, Sheldon O.
 ; REGISTRATION NUMBER: 38,179
 ; REFERENCE/DOCKET NUMBER: 19846NP2
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 732-594-1958
 ; TELEFAX: 732-594-4720
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 283 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: Other
 ; FEATURE:
 ; NAME/KEY: Other
 ; LOCATION: 1...283
 ; OTHER INFORMATION: cdna probe

US-08-993-424B-4
 Query Match 20.4%; Score 283; DB 4; Length 283;
 Best Local Similarity 100.0%; Pred. No. 2.6e-48;
 Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 206 TGGGACCAACCACTTGTACCTGGGAGCATGGCCGTGTCGACCTACTATCTGTC 265
 Db 1 TGGGACCAACCACTTGTACCTGGGAGCATGGCCGTGTCGACCTACTATCTGTC 60

QY 266 TCGGGCTGCGGTTGACCTGTACCGCTCTGGGGCTCGCGCCCTGGGGCGC 325
 Db 61 TCGGGCTGCGGTTGACCTGTACCGCTCTGGGGCTCGCGCCCTGGGGCGC 120

QY 326 TGCTCTCGCGCTGTCCCTCTACGTGGGCGAGGCTGCACCTAGCCACGCTGTGCACA 385
 Db 121 TGCTCTCGCGCTGTCCCTCTACGTGGGCGAGGCTGCACCTAGCCACGCTGTGCACA 180

QY 386 TGACCGGCTCAGGCTCGAGGCTACCTGGCCATCTGCCGCCGCTCGCGCGCGGTCT 445
 Db 181 TGACCGGCTCAGGCTCGAGGCTACCTGGCCATCTGCCGCCGCTCGCGCGCGGTCT 240

QY 446 TGGTACCCGCGCGCGCTCGCGCGCTCATCGCTGTGCTCTG 488
 Db 241 TGGTACCCGCGCGCGCTCGCGCGCTCATCGCTGTGCTCTG 283

RESULT 3
 US-09-603-680-4
 ; Sequence 4, Application US/09603680
 ; Patent No. 6544753
 ; GENERAL INFORMATION:
 ; APPLICANT: Tan, Carina
 ; APPLICANT: Sullivan, Kathleen
 ; TITLE OF INVENTION: GALANIN RECEPTOR GALR2 AND
 ; NUCLEOTIDES ENCODING SAME
 ; NUMBER OF SEQUENCES: 20
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Merck & Co., Inc.
 ; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
 ; CITY: Rahway
 ; STATE: NJ
 ; COUNTRY: USA
 ; ZIP: 07065-0900
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: Windows
 ; SOFTWARE: FASTSEQ For Windows Version 2.0b
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/603,680
 ; FILING DATE: 26-Jun-2000
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/033,851
 ; FILING DATE: 27-DEC-1996
 ; APPLICATION NUMBER: 08/993,088
 ; FILING DATE: 18-DEC-1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Heber, Sheldon O.
 ; REGISTRATION NUMBER: 38,179
 ; REFERENCE/DOCKET NUMBER: 19846 CA
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 732-594-1958
 ; TELEFAX: 732-594-4720
 ; TELEX: <Unknown>
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 283 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: Other


```

FEATURE:
NAME/KEY: Other
LOCATION: 1...283
OTHER INFORMATION: cDNA probe
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-603-680-4

Query Match
Best Local Similarity 20.4%; Score 283; DB 4; Length 283;
Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 206 TGGGACCCACCACTTCTACTGCGGAGCATGGCGTTCGACCTACTCTCTGC 265
DB 1 TGGGACCCACCACTTCTACTGCGGAGCATGGCGTTCGACCTACTCTCTGC 60

QY 266 TCGGGGCTGCCGTTTCGACCTGTACCGCTCTGGCGCTCGCGGCTTGGGCGCG 325
DB 61 TCGGGGCTGCCGTTTCGACCTGTACCGCTCTGGCGCTCGCGGCTTGGGCGCG 120

QY 326 TGCTTGCCTGCTGCTCTTACGTGGGCGAGGCTGCACCTACGCCACGCTCTGCACA 385
DB 121 TGCTTGCCTGCTGCTCTTACGTGGGCGAGGCTGCACCTACGCCACGCTCTGCACA 180

QY 386 TGACCGGCTCAGGCTCGAGGCTTACGTGGGCGAGGCTGCACCTACGCCACGCTCTGCACA 445
DB 181 TGACCGGCTCAGGCTCGAGGCTTACGTGGGCGAGGCTGCACCTACGCCACGCTCTGCACA 240

QY 446 TGCTTACCGGCGCGGCTCGCGGCTCATCGTGTGCTCTG 488
DB 241 TGCTTACCGGCGCGGCTCGCGGCTCATCGTGTGCTCTG 283

```

```

RESULT 4
US-09-77-675A-1
Sequence 1, Application US/09077675A
Patent No. 6242199
GENERAL INFORMATION:
APPLICANT: Pai, Lee-Yuh
APPLICANT: Feighner, Scott C.
APPLICANT: Howard, Andrew D.
APPLICANT: Pong, Sheng-Shung
APPLICANT: Van Der Ploeg, Leonardus H.T.
TITLE OF INVENTION: RECEPTOR ASSAY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION NUMBER: US/09/077,675A
FILING DATE: 3-JUN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19590P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 1063 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-077-675A-1

Query Match
Best Local Similarity 16.9%; Score 234.8; DB 3; Length 1063;
Matches 320; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 92 CGCCCTTTCCCTTGGGGGCGCTGGTCCCGTACCGCTGTGTCCTGTCGCTGCTCGTCG 151
DB 69 CGCTCTTCCCTTGGGGGCGCTGGTCCCGTACCGCTGTGTCCTGTCGCTGCTCGTCG 128

QY 152 TCGGGGTGAGCGGCAACGTTGGTACCGTGTATGTCGCGGCGCTACCGGACATGCGCA 211
DB 129 TGGGTATCGGGGCAACCTGCTCACGATGCTGGTAGTGTACGCTTCCGGGAGATGGCA 188

QY 212 CCACCAACCACTTGTACTTGGGAGCATGGCGGTGTCGACCTACTCTCTCTCGGGC 271
DB 189 CCACCAACCACTTGTACTTGGGAGCATGGCGGTGTCGACCTACTCTCTCTCTGCA 248

QY 272 TCGCGTTTCACCTGTACCGCTCTGGGGCTCGGGGCTGGGGTGTGGGCGCGCTGCTCT 331
DB 249 TCGCGTTTCACCTGTACCGCTCTGGGGCTCGGGGCTGGGGTGTGGGCGCGCTGCTCT 308

QY 332 GCGCGCTGCTCCCTTACGTGGGAGGCTGACCTACGCGCGCTGTCACATGACCG 391
DB 309 GCAACTCTTCAGTTCGTTAGGAGAGTGCACCTACGCGCAGTGTCTACCATCACC 368

QY 392 CGCTACGCGTCGAGCGCTACCTGCGCATCTGCGCGCGCTGCGCGCGCGCTTGTGTCA 451
DB 369 CGCTAGCGTCGAGCGCTACTTCCGCATCTGCTTCCCGCTGCGGCGCAAGGTAGTCA 428

QY 452 CCGCGCGCGCGCTGCGCGCGCTCATCGCTGTGCTCTGGGCGGTGGGCGCTCTCTGCGCG 511
DB 429 CCAAGGCGCGGTAAGCTGCTCATCTGGTGTATCTGGGCGGTGGGCTTCTGCGAGCGCG 488

QY 512 GTCCCTTTCTTGTTCCTGTTGGGCGTCGAGCAGGACCGCGCA 553
DB 489 GGCCCATCTTCGTGCTGGTGGAGTGAGCATGATATACGCA 530

```

```

RESULT 5
US-09-077-674-1
Sequence 1, Application US/09077674
Patent No. 6531314
GENERAL INFORMATION:
APPLICANT: Arena, Joseph P.
APPLICANT: Cully, Doris P.
APPLICANT: Feighner, Scott D.
APPLICANT: Howard, Andrew D.
APPLICANT: Liberator, Paul A.
APPLICANT: Schaeffer, James M.
APPLICANT: Van Der Ploeg, Leonardus
TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR FAMILY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/077,674
FILING DATE: 3-JUN-1998

```

```
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19589P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1063 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-09-719-674-1

Query Match 16.9%; Score 234.8; DB 4; Length 1063;
Best Local Similarity 69.3%; Pred. No. 1.4e-38;
Matches 320; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 92 CGCCCTTTCCCTGGGGCGCTGGTCCGGTGACCGCTGTGTGCTGTGCTGTGCTGTGCTG 151
DB 69 CGCTCTTCCACGCGCTGTGGGGCGCTACCGGCACCTGGTGGCGCTCTTGTGG 128
QY 152 TCGGGGTGACGGCAACGCTGTGACCGTGATGTGTGATCGGGCGCTACCGGGACATGCGGA 211
DB 129 TGGGTATCGGGGCAACCTGCTCAGGATGCTGGTGTGATGTGTGATGTGTGATGTGTG 188
QY 212 CCACCAACCACTTGTACCTGGGAGCATGCGGTGTGCTGTGCTGTGCTGTGCTGTGCTG 271
DB 189 CCACCAACCACTTGTACCTGGGAGCATGCGGTGTGCTGTGCTGTGCTGTGCTGTGCTG 248
QY 272 TGGGTATCGGGGCAACCTGCTCAGGATGCTGGTGTGATGTGTGATGTGTGATGTGTG 331
DB 249 TGGGTATCGGGGCAACCTGCTCAGGATGCTGGTGTGATGTGTGATGTGTGATGTGTG 308
QY 332 GCGCCCTGTCCCTTACGTGGGGGAGGGGTGCACTACGCAACGCTGTGCTGTGCTGTGCTG 391
DB 309 GCAAACTCTTCCAGTTCGTTAGCGAGAGCTGCACCTACGCAACAGTGTCTCACCATCACCG 368
QY 392 CGCTCAGCTCGAGCGCTACTGGCCATCTGCGCGCGCTTCCGGCGCGCGCTTCTGTGCTCA 451
DB 369 CGCTGAGCTCGAGCGCTACTTCCGCATCTGCTTCCGGCGCGCGCTTCTGTGCTCA 428
QY 452 CCGCGCGCGCGCTCGCGCGCTCATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 511
DB 429 CCAGGGCCGGGTAAAGCTGGTATCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 488
QY 512 GTCCCTTCTTGTTCCTGGTGGGGTTCGAGCAGGACCCCGGCA 553
DB 489 GGCCCATCTTCGTGCTGGTGGGAGTGAGCATGATAACGGCA 530

RESULT 6
US-09-719-675A-4
; Sequence 4, Application US/09077675A
; Patent No. 6242199
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
```

```
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/077,675A
FILING DATE: 3-JUN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19590P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1029 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-09-077-675A-4

Query Match 16.5%; Score 230; DB 3; Length 1029;
Best Local Similarity 68.8%; Pred. No. 1.3e-37;
Matches 317; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 92 CGCCCTTTCCCTGGGGCGCTGGTCCGGTGACCGCTGTGTGCTGTGCTGTGCTGTGCTG 151
DB 266 CGCTCTTCCACGCGCTGTGGGGCGCTACCGGCACCTGGTGGCGCTCTTGTGG 325
QY 152 TCGGGGTGACGGCAACGCTGTGTGACCGTGATGTGTGATCGGGCGCTACCGGGACATGCGGA 211
DB 326 TGGGTATCGGGGCAACCTGCTCAGGATGCTGGTGTGATGTGTGATGTGTGATGTGTG 385
QY 212 CCACCAACCACTTGTACCTGGGAGCATGCGGTGTGCTGTGCTGTGCTGTGCTGTGCTG 271
DB 386 CCACCAACCACTTGTACCTGGGAGCATGCGGTGTGCTGTGCTGTGCTGTGCTGTGCTG 445
QY 272 TGGGTATCGGGGCAACCTGCTCAGGATGCTGGTGTGATGTGTGATGTGTGATGTGTG 331
DB 446 TGGGTATCGGGGCAACCTGCTCAGGATGCTGGTGTGATGTGTGATGTGTGATGTGTG 505
QY 332 GCGCCCTGTCCCTTACGTGGGGGAGGGGTGCACTACGCAACGCTGTGCTGTGCTGTGCTG 391
DB 506 GCAAACTCTTCCAGTTCGTTAGCGAGAGCTGCACCTACGCAACAGTGTCTCACCATCACCG 565
QY 392 CGCTCAGCTCGAGCGCTACTGGCCATCTGCGCGCGCTTCCGGCGCGCGCTTCTGTGCTCA 451
DB 566 CGCTGAGCTCGAGCGCTACTTCCGCATCTGCTTCCGGCGCGCGCTTCTGTGCTCA 625
QY 452 CCGCGCGCGCGCTCGCGCGCTCATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 511
DB 626 CCAGGGCCGGGTAAAGCTGGTATCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 685
QY 512 GTCCCTTCTTGTTCCTGGTGGGGTTCGAGCAGGACCCCGGCA 553
DB 686 GGCCCATCTTCGTGCTGGTGGGAGTGAGCATGATAACGGCA 727

RESULT 7
US-09-077-674-4
; Sequence 4, Application US/09077674
```



```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19589P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1122 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-077-674-9

Query Match          16.5%; Score 228.8; DB 4; Length 1122;
Best Local Similarity 57.9%; Pred. No. 2.3e-37;
Matches 424; Conservative 0; Mismatches 307; Indels 1; Gaps 1;

QY   93 GCCTTTTCCCTGGGCGCGCTGTGTGCCGTGACCGCTGTGCCTGTGCCTGTGCCTTCGTCTG 152
DB   360 GCTCTTCCC CGCGCGCTGTGTGGCGGCGCTCACAGCCACCCTGCGTGCGACTCTTCTGCGGT 419

QY   153 CGGGGTGAGCGGCAACTGTGTGACCGGTATGCTCATCGGCGGCTACCGGGACATGCGGAC 212
DB   420 GGGTATCGCTGGCAAACTGCTCACCATGCTGTGTGGTGTGGCGCTTCGCGAGCTGCCGAC 479

QY   213 CACCACCAACTGTGTACCTGGGCAGCATGGCGGTGTCCGACCTCATCTCTGCTCGGGCT 272
DB   480 CACCACCAACTCTACCTGTGTCCAGCATGGCTTCTCGAATGCTGCTCATCTTCTCTCGAT 539

QY   273 GCGGTTGCACTGTACCGCTCTGTGGCGCTGTGGCGCCCTGTGGGTGTTCGGGCGCGTGTCTG 332
DB   540 GCCCTTGAGACTCGTTTCGCCCTCTGGCAGTAGCGGCCCTTGGAACTTCGCGACCTCTCTG 599

QY   333 CCGGCTGTCCCTCTACCTGTGGCGAGGCTGCACCTAGCCACGCTGTGCACATGACCGC 392
DB   600 CAAACTCTTCAAATTCGTGAGTGAAGCTGCACCTAGCCACGGTGCTCACCATCA CAGC 659

QY   393 GCTCAGCGTTCGAGCGCTACCTGGCCATCTGCGGCCGCTTCGCGCCCGCGCTTTGGTCA 452
DB   660 GCTGAGCGTCGAGCGCTACTTGGCCATCTGCTTCCACTTCGGGCCAAGTGTGGTCA 719

QY   453 CCGGCGCGCGGTTCGCGCGCTCATCGTGTGCTCTGGGCGGTGTGGCGGTGCTCTCTGCGCG 512
DB   720 CAAGGGCGGGGTGAAGCTGTGTCATCTTCTGTCATCTGGGCGGTGGCCCTTCTGCAGCGCGG 779

QY   513 TCCCTTCTGTCTGTGTGGGCGTCGAGCAGGACCCCGGCATCT-CCGTAGTCCCGGGCC 571
DB   780 GCCCATCTTGTGTGTGCGGGTGGAGCAGAGAACCGCACCCACCTTTGGGACACCAA 839

QY   572 TCAATGCAACCGCGCGATCGCCTCTCGCTCTGCGCTCTGTCCCGCTCTCTGGGTCT 631
DB   840 CGAGTGGCGCCCAACGAGTTTGGGTGGCTCTGGACTGCTCACGCTCATGTTGGT 899

QY   632 CGGGGCGCCACCGCGCTTCCCGCGGTGGGGCCCCGAGACGCGGAGGCGCGCGCGCTGT 691
DB   900 GTCCAGCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 959

QY   692 TCAGCCGGAATGCCGCGGNGCCCCCGGAGCTGGGGCGGCTGTGTCATGCTGTGGG 751
DB   960 CAGGAAGCTGTGGCGGAGAGGCGCGCGCATGTGTGTCTGTGGTGTCCCTCGCTCAGGGACA 1019

QY   752 TCACACCGGCTACTTCTTCTGCGCCCTTTCTGTGCTTCTGCTCAGCATCTCTACGGGCTCATCG 811
DB   1020 GNACCAAGAACAACCGTGAAAATGCTGGGTGGTCTCAGCGCGCGCTCAGGCTTCTCT 1079

QY   812 GCGCGGAGCTGT 823
```

```

Db      1080  CGCGGGTCCTAT 1091
||||| |||
RESULT 11
US-09-077-675A-15
; Sequence 15, Application US/09077675A
; Patent No. 6242199
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1092 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-077-675A-15

Query March 16.4%; Score 228.4; DB
Best Local Similarity 68.4%; Pred.No.2.7e-3
Matches 316; Conservative 0; Mismatches 1

Qy      92  CGCCCTTTCCCTCGGGGCGGTGGTGGCGCGGTGACCC
Db      104  GCCTGTTCCCGCGCTCCGCTGTGGCAGGCGTCACCG
Qy      152  TCGGGGTGAGCGGCAACGTGGTGACCGTGAATGCTGA
Db      164  TGGGCATCTCAGGCAACCTGTCTCACTATGCTGGTGG
Qy      212  CCACCAACAATTGTACTCTGGCAGCATGGCCGTGT
Db      224  CCACCAACAACCTCTACCTGTCCAGCATGGCCCTTCT
Qy      272  TGGCGTTGCACCTGTACCGCCTCTGGCGGTCCGCGG
Db      284  TGCCGCTGGACCTCGTCGCCTCTGGCAGTACCGG
Qy      332  GCCGCCTGTCCCTCTTACCTGGGCGAGGGCTGCACCT

```



```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
US-09-077-675A-14

Query Match      16.4%; Score 228.4; DB 3; Length 3129;
Best Local Similarity 68.4%; Pred. No. 3:2e-37;
Matches 316; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

Qy   92 CGCCCTTTCCCTGGGCGCGTGTGCCTGGTGACCGCTGTGTGCTGTGCTGTTCGTTCGTCG 151
Db   104 CGCTGTTCCCGCTCCGCTGCTGGCAGCGCTCACCGCACCTGGTGGCGCTCTTCGTGG 163
Qy   152 TCGGGGTGAGCGGCAACGTGTGACCCTGATCTGTATCGGCGCTACCGGACCATGCGGA 211
Db   164 TGGGCATCTCAGGCCAACCTGCTCACTATGCTGGTGGTGTCCCGCTTCGCGAGCTGCGCA 223
Qy   212 CCACCCAACTTCTACTCTGGGAGCATGGCGGTGTCGACCTACTCATCTGCTCGGGC 271
Db   224 CCACCCAACTTCTACTCTGCCAGCATGGCGCTTCTCGGATCTGCTCATCTTCTGTGCA 283
Qy   272 TGCGGTTTCGACCTGTACCGCTCTGGCGCTCGCGGCCCTGGGTGTTTCGGGCGGTGCTCT 331
Db   284 TGCGCTTGGACCTCGTCCGCTCTGGCAGTACCGGCCCTGGAACTTCGGCGACCTGCTCT 343
Qy   332 GCOCCTGTCCCTCTACGTGGGCGAGGCTGCACCTACGCCACCGTGTGCACATGACCG 391
Db   344 GCAAATCTTCCAGTTTGTGAGGAGAGCTGCACCTACGCCACCGTCTCACCATCACCG 403
Qy   392 CGCTCAGGCTCGAGCGCTACCTGGCCATCTCGCGCCCGCTCGGGCCCGCGCTTTGGTCA 451
Db   404 CGCTGAGCGTGCAGCGCTACTTCCGCATCTGCTTCCCTCTCGGGGCCAAGGTGGTGTA 463
Qy   452 CCGGGCCCGGTCGCGCGCTCATCGCTGTGCTGTGGGCGTGGCGCTGCTCTCTGCGCG 511
Db   464 CTAAAGGCGCGGTGAAGCTGTCTATCTTGTTCATCTGGGCGCGTGGCTTCTGACGCGCG 523
Qy   512 GTCCCTTTTGTTCCTGTGGGCGTTCGAGCAGGACCCCGGCA 553
Db   524 GGCCCATCTTCGTGCTGGTGGGCGTGGAGCAGCAAAAACGGCA 565

RESULT 14
US-09-077-674-14
; Sequence 14, Application US/09077674
; Patent No. 6531314
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR FAMILY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,674
; FILING DATE: 3-JUN-1998
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
```


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Db 421 TGCGCCCGCTCCGCGCCCGCTTGGTACCCGCGCGCGCTCCGCGCTCATCGCT 480
Qy 481 GTGCTCTGGGCGCGCTGCTCTCTGCGCGCTCCCTTCTTGTCTGCTGGGCGTGGAG 540
Db 481 GTGCTCTGGGCGCGCTGCTCTCTGCGCGCTCCCTTCTTGTCTGCTGGGCGTGGAG 540
Qy 541 CAGACCCCGGATCTCCGCTAGTCCCGGCTCAATGCGACCGCGCGATCGCTCTCG 600
Db 541 CAGACCCCGGATCTCCGCTAGTCCCGGCTCAATGCGACCGCGCGATCGCTCTCG 600
Qy 601 CCTCTCGCTCTGCGCGCTCTCTGCTCTGCGCGCGCGCTCCGCGCTCCGCGCTCG 660
Db 601 CCTCTCGCTCTGCGCGCTCTCTGCTCTGCGCGCGCGCTCCGCGCTCCGCGCTCG 660
Qy 661 GGGCCCGAGACCGCGGAGCGCGCGCTGTTACGCGCGGAATCCGCGCGAGCCCGCG 720
Db 661 GGGCCCGAGACCGCGGAGCGCGCGCTGTTACGCGCGGAATCCGCGCGAGCCCGCG 720
Qy 721 CAGTGGGCGCTGCTGCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 721 CAGTGGGCGCTGCTGCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Qy 781 CTGTGCTCAGATCTCTACGGCTCATCGGGCGGAGCTGTGGAGACCGCGCGCG 840
Db 781 CTGTGCTCAGATCTCTACGGCTCATCGGGCGGAGCTGTGGAGACCGCGCGCG 840
Qy 841 CTGCGAGCGCGCGCGCTCGGGCGGAGAGCGCCACCGGAGAGCGCTCGCGCTG 900
Db 841 CTGCGAGCGCGCGCGCTCGGGCGGAGAGCGCCACCGGAGAGCGCTCGCGCTG 900
Qy 901 CGTAAGTGGAGCGCGCGTGTTCAGAGAGCGCTGCTGCGAGCTCGCGCGCGGAGC 960
Db 901 C----- 901
Qy 961 GCGCAAGCTGGGTCCCTTCCCTGCTGCCCGAGCTCTGGGCGCGCTTCCAGTCCC 1020
Db 902 ----- 902
Qy 1021 TTTCTATTTCGATTCAGCTCCACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Db 902 -----TGGTGGTGGTCTTGGCAITTAATTTG 929
Qy 1081 CTGTTGCTCTTCCAGCTTGGCAGATCAATTAATAAACAAGAGATTTCGCGATGAT 1140
Db 930 CTGTTGCTCTTCCAGCTTGGCAGATCAATTAATAAACAAGAGATTTCGCGATGAT 989
Qy 1141 GTACTTCTCAGTACTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
Db 990 GTACTTCTCAGTACTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1049
Qy 1201 CAACCCAACTCTTCAACCTTCAAGAGATCAGAGCGCGCGCTTAAACTGCT 1260
Db 1050 CAACCCAACTCTTCAACCTTCAAGAGATCAGAGCGCGCGCTTAAACTGCT 1109
Qy 1261 GCTCGAAGAGTCCAGCGCGAGAGCTTCCAGAGAGAGAGAGCTGCGGGGAGT 1320
Db 1110 GCTCGAAGAGTCCAGCGCGAGAGCTTCCAGAGAGAGAGAGCTGCGGGGAGT 1169
Qy 1321 TGAGGGGACACTGGAGAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
Db 1170 TGAGGGGACACTGGAGAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1229
Qy 1381 GATGGGATAA 1390
Db 1230 GATGGGATAA 1239

RESULT 3
US-10-290-078-13
; Sequence 13, Application US/10290078
; Publication No. US20030124596A1
; GENERAL INFORMATION:

; APPLICANT: Carroll, Joseph A.
; TITLE OF INVENTION: Hematological Disorders Using 232, 2059, 10630, 12848, 13875,
; FILE OF INVENTION: 14395, 14618, 17692 or 58874
; FILE REFERENCE: MPI2001-288P1(M)
; CURRENT APPLICATION NUMBER: US/10/290,078
; CURRENT FILING DATE: 2002-11-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 1239
; TYPE: DNA
; ORGANISM: Homo Sapien
; US-10-290-078-13

Query Match 77.6%; Score 1078; DB 15; Length 1239;
Best Local Similarity 89.1%; Pred. No. 5.8e-255;
Matches 1239; Conservative 0; Mismatches 0; Indels 151; Gaps 1;
Qy 1 ATGGGCGAGCCCTGGAAACGCGACGCGCCCGAGGGGCGCGGAGCGCGTGGCCCC 60
Db 1 ATGGGCGAGCCCTGGAAACGCGACGCGCCCGAGGGGCGCGGAGCGCGTGGCCCC 60
Qy 61 GCCTGCGCGCTTGGACGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
Db 61 GCCTGCGCGCTTGGACGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
Qy 121 GTACACGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
Db 121 GTACACGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
Qy 181 ATGCTGATCGGGCGCTACCGGGACATGCGGACACCAACCACTTGTACCTGGGCGAGCATG 240
Db 181 ATGCTGATCGGGCGCTACCGGGACATGCGGACACCAACCACTTGTACCTGGGCGAGCATG 240
Qy 241 GCGGTGTGCGACCTACTCATCTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 300
Db 241 GCGGTGTGCGACCTACTCATCTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 300
Qy 301 TCGCGCGCTGCGGCTGCTGCGGCGCTGCTGCGGCGCTGCTGCGGCGCTGCTGCGGCGCTGCT 360
Db 301 TCGCGCGCTGCGGCTGCTGCGGCGCTGCTGCGGCGCTGCTGCGGCGCTGCTGCGGCGCTGCT 360
Qy 361 TGACACCTACGCCACGCTGCTGACATGACCGCGCTCAGCGCTCAGCGCTCAGCGCTACCTGGCCATC 420
Db 361 TGACACCTACGCCACGCTGCTGACATGACCGCGCTCAGCGCTCAGCGCTACCTGGCCATC 420
Qy 421 TGCGCGCGCTTCCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Db 421 TGCGCGCGCTTCCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Qy 481 GTGCTCTGGGCGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Db 481 GTGCTCTGGGCGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Qy 541 CAGACCCCGGATCTCCGCTAGTCCCGGCTCAATGGACACCGCGGATCGCTCTCG 600
Db 541 CAGACCCCGGATCTCCGCTAGTCCCGGCTCAATGGACACCGCGGATCGCTCTCG 600
Qy 601 CCTCTGCTGCTGCGCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 601 CCTCTGCTGCTGCGCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Qy 661 GGGCCCGAGACCGCGGAGCGCGCGCTGTTACGCGCGGAATGCGCGCGAGCCCGCG 720
Db 661 GGGCCCGAGACCGCGGAGCGCGCGCTGTTACGCGCGGAATGCGCGCGAGCCCGCG 720
Qy 721 CAGTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 721 CAGTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Qy 781 CTGTGCTCAGATCTCTTACGGCTCATCGGGCGGAGCTGTGGAGACCGCGCGCG 840

213 CACACCAAACTTGTACCTGGGAGCATGGCGGTGTCGACCTACTCATCTCTGCTCGGGCT 272
| | | | |
Db CACACCAAACTTGTACCTGGGAGCATGGCGGTGTCGACCTACTCATCTCTGCTCGGGCT 287
| | | | |
QY GCGTTCGACCTGTACCGCTTGTGGCGCTCGCGCCCTGGGTGTTGGGCGCTGCTCTG 332
| | | | |
Db GCGCTTCGACCTGTACCGCTTGTGGCGCTCGCGCCCTGGGTGTTGGGCGCTGCTCTG 347
| | | | |
QY CCGCTTGTCTCTTACGTGGGCGAGGCTGACCTAGCGCAGCATGCTGTCACATGACCGC 392
| | | | |
Db CAAACTCTTCAATGCTAGTGAGAGCTGACCTAGCGCAGCATGCTGTCACATGACAGC 407
| | | | |
QY GCTCAGCGTTCAGCGCTTACCTGGCCATCTGCGCCCGCTCGCGCCCGCTGCTTGGTTCAC 452
| | | | |
Db GCTCAGCGTTCAGCGCTTACCTGGCCATCTGCGCCCGCTCGCGCCCGCTGCTTGGTTCAC 467
| | | | |
QY CCGCGCGCGCTGCGCGCTCATGCTGTGCTGTGGCGCGTGGCGCTGCTCTCTGCGCG 512
| | | | |
Db CAAAGGGCGGGTGAAGCTGCTCATCTTCTGTCATCTGGGCGTGGCGCTTCTGACAGCGCGG 527
| | | | |
QY TCCCTTCTTCTGCTGCTGGGCGTTCGAGCAGGACCGCGGCTCT - CCGTAGTCCCGGGC 571
| | | | |
Db GCCCATCTTCTGCTAGTGGGCTGGAGCAGAGACCGGACCGACCCCTTGGGACACCAA 587
| | | | |
QY TCAATGGCACCGCGGATCGCTCTCTGCGCTCTGCGCTCGTCCGCGCTCTCTGCGTCT 631
| | | | |
Db CGAGTGGCGGCGGACCGAGTTGGGCTGCGCTCTGAGCTGCTCAGCGTCTGCTGGGT 647
| | | | |
QY CCGGGGGCGCACCGCGCTTCCCGCGTGGGCGGCGGAGACCGCGGAGCGCGCGGCGTGT 691
| | | | |
Db GTCCAGCATCTTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 707
| | | | |
QY TCAGCGCGGAATCGCGGCGGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 751
| | | | |
Db CAGAAAGCTGTGGCGGAGGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 767
| | | | |
QY TCACCAACCGCTTCTTCTTCTGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 811
| | | | |
Db GAACCAACGAAACCGTGAATGCTGGGTGGTCTCAGCGCGCTCAGGGCTTCTCT 827
| | | | |
QY GCGCGGAGCTGT 823
| | | | |
Db CCGCGGCTCTAT 839

RESULT 10

US-10-303-204A-9
; Sequence 9, Application US/10303204A
; Publication No. US20030166144A1
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus H. T.
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR
; FILE REFERENCE: 19589PCA
; CURRENT APPLICATION NUMBER: US/10/303,204A
; PRIOR FILING DATE: 2002-11-25
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: PCT/US96/19445
; PRIOR FILING DATE: 1996-12-10
; PRIOR APPLICATION NUMBER: 60/018,962
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: 60/008,582
; PRIOR FILING DATE: 1995-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9

; LENGTH: 1122
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-303-204A-9

Query Match 16.5%; Score 228.8; DB 13; Length 1122;
Best Local Similarity 57.9%; Pred. No. 1.7e-46;
Matches 424; Conservative 0; Mismatches 307; Indels 1; Gaps 1;

QY 93 GCGCTTCCCTGGGCGCTGTCGCGGTGACCGCTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 132
| | | | |
Db 360 GCTCTTCCCGCGCGCTGTCGCGGCGTTCACAGCCACCTGCGTGGCACTCTTCTGCTGCTG 419
| | | | |
QY 153 CCGGGTGAGCGGCAACGCTGATGACCTGATGATGATGATGATGATGATGATGATGATGATGATG 212
| | | | |
Db 420 GGGTATCGTGCACCACTGCTCACCATGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 479
| | | | |
QY 213 CACCAACAACTTGTATCTGGGAGCATGATGCGGTGTCGACCTACTACTACTACTACTACTACTACT 272
| | | | |
Db 480 CACCAACAACTTGTATCTGGGAGCATGATGCGGTGTCGACCTACTACTACTACTACTACTACTACT 539
| | | | |
QY 273 GCGGTCGACCTGTCGCGCTCTGCGGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCTG 332
| | | | |
Db 540 GCGGTCGACCTGTCGCGCTCTGCGGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCTG 599
| | | | |
QY 333 CCGCTGTCCTCTTACGTGGGCGAGGCTGACCTAGCGCAGCATGCTGCTGCTGCTGCTGCTGCTG 392
| | | | |
Db 600 CAAACTCTTCCAACTTCTGCTAGTGAGAGCTGACCTAGCGCAGCATGCTGCTGCTGCTGCTG 659
| | | | |
QY 393 GCTCAGCGTTCAGCGCTTACCTGGCCATCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCTGCTG 452
| | | | |
Db 660 GCTCAGCGTTCAGCGCTTACCTGGCCATCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCTGCTG 719
| | | | |
QY 453 CCGGCGCGCGTTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 512
| | | | |
Db 720 CAAAGGGCGGGTGAAGCTGCTCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 779
| | | | |
QY 513 TCCCTTCTTCTTCTGCTGGTGGGCTGAGCAGGACCGCGGCTGCT - CCGTAGTCCCGGGC 571
| | | | |
Db 780 GCCCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 839
| | | | |
QY 572 TCAATGGCACCGCGGATCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 631
| | | | |
Db 840 CGAGTGGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 899
| | | | |
QY 632 CCGGCGGCGCACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 691
| | | | |
Db 900 GTCCAGCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 959
| | | | |
QY 692 TCAGCGCGGAATGCGCGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 751
| | | | |
Db 960 CAGGAAGCTGTGGCGGAGGAGCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1019
| | | | |
QY 752 TCACCAACCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 811
| | | | |
Db 1020 GAACCAACGAAACCGTGAATGCTGGGTGGTCTCAGCGCGCTCAGGGCTTCTCTCTCTCTCT 1079
| | | | |
QY 812 GCGCGGAGCTGT 823
| | | | |
Db 1080 CCGCGGCTCTAT 1091

RESULT 11
US-10-303-204A-15
; Sequence 15, Application US/10303204A
; Publication No. US20030166144A1
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.

Qy	153	CGGGTGAGCGGCAACGTGTGACCGTGATGCTGATCGGGCGCTACCGGACATGGGAC	212
Db	168	GGGTATCGTGGCAACCTGCTACCAATGCTGGTGGTGTGCGGCTTCGCGAGCTGGGCAC	227
Qy	213	CACCACCAACTTGTACCTGGGAGCATGGCGGTTCGACCTACTCATCTCTCGGGCT	272
Db	228	CACCACCAACTCTACCTGTCCAGCATGGCTTCTCCGATCTGCTCATCTTCTCTGCAT	287
Qy	273	GCGTTGACCTGTACCGCCTCTGGCGCTCGCGCCCTGGGTGTTGGGGCCGCTGCTCTG	332
Db	288	GCCCTGGACCTCGTTGCGCTCTGGCAGTACCGGCCCTGGAACTTCGGCGACCTCCTCTG	347
Qy	333	CGGCTGTCCCTCTACGTGGGAGGGCTGGACCTAGCCACGGCTGTGCACATGACCGC	392
Db	348	CAAACTCTTCCAATTCTGTCAGTGAGAGCTGCACCTACGCCACGGTGTCAACCATCACAGC	407
Qy	393	GCTCAGGTCGAGCGCTACCTGGCCATCTCGCGCCGCTCGCGCCGCGTCTTGGTCAC	452
Db	408	GCTGAGGCTCGAGCGCTACTTCGCCATCTGTCCCACTCCGGGCCAAGGTGGTGGTCAC	467
Qy	453	CCGGCGCGCGTCCGCGCGCTCATCGCTGTGCTCTGGGCGGTGGCGCTGCTCTGCGCG	512
Db	468	CAAGGGCGGGTGAAGCTGGTCATCTTCGTCACTCTGGCGCGTGGCCTTCTGCAGCGCGG	527
Qy	513	TCCCTTCTTGTTCCTGGTGGCGTTCGAGCAGGACCCCGGCATC	555
Db	528	GCCCATCTTCGTGTAGTCGGGGTGGAGCACGAGACGGCACC	570

Search completed: January 1, 2004, 06:44:59
Job time : 693.92 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 31, 2003, 21:38:58 ; Search time 3259.02 Seconds
(without alignments)
10366.072 Million cell updates/sec

Title: US-09-719-485-4

Perfect score: 1390

Sequence: 1 atgggcagccctggaacgg.....acgtgaagacgatgggataa 1390

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_estl:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_man:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	230	16.5	4435	11 AK049671	AK049671 Mus muscu
2	226.8	16.3	500	10 BF603623	BF603623 269181 MA
3	219.2	15.8	608	14 BY724644	BY724644 BY724644
4	218	15.7	590	10 BF513101	BF513101 UI-H-BW1-

5	186.4	13.4	843	13	BUS553576	BUS553576
C	6	173.4	12.5	849	13	BUS68940
C	7	141.2	10.2	504	28	AZ451922
8	127	9.1	855	12	BI757121	BI757121
9	127	9.1	1010	12	BM925480	BM925480
10	125.4	9.0	634	10	BG714306	BG714306
11	125.4	9.0	1055	12	BM808829	BM808829
12	125.2	9.0	540	12	BM737726	BM737726
13	124.8	9.0	1027	12	BM543497	BM543497
14	124	8.9	1144	12	BM546496	BM546496
C	15	122.8	8.8	941	29	CNS036WH
16	119	8.6	1603	12	BM808823	BM808823
17	117.8	8.5	624	9	AU244292	AU244292
18	111.8	8.0	738	12	BI600170	BI600170
19	111.4	8.0	426	14	CB695276	CB695276
20	111.4	8.0	574	14	CB608597	CB608597
21	111.2	8.0	928	12	BI758259	BI758259
22	109.6	7.9	993	13	BQ721315	BQ721315
23	109.2	7.9	448	10	BF323227	BF323227
24	106.4	7.7	3277	11	AK036756	AK036756
25	105.6	7.6	2760	11	AK081073	AK081073
26	105	7.6	427	13	BY284735	BY284735
27	104.8	7.5	828	12	BI597845	BI597845
28	104.6	7.5	660	14	BY728041	BY728041
29	103.6	7.5	866	14	CD246184	CD246184
30	102.6	7.4	980	12	BM543468	BM543468
31	102.4	7.4	419	13	BY273314	BY273314
32	102.2	7.4	994	14	BY705540	BY705540
33	102.2	7.4	1189	11	AK005368	AK005368
C	34	99.4	7.2	1233	14	CA975828
35	99	7.1	397	14	CB772331	CB772331
36	98.8	7.1	1918	11	AK053776	AK053776
37	98.4	7.1	751	13	CB154463	CB154463
C	38	97.8	7.0	982	13	BX415111
39	97	7.0	785	12	BI754749	BI754749
40	97	7.0	843	29	CNS046SGJ	CNS046SGJ
41	96.8	7.0	670	29	AG046172	AG046172
42	96.8	7.0	720	12	BI753905	BI753905
43	96.8	7.0	1307	11	CNSLT11BD	CNSLT11BD
44	96.6	6.9	555	10	BE751626	BE751626
45	96.6	6.9	825	9	AU079556	AU079556

ALIGNMENTS

RESULT 1

AK049671

LOCUS

DEFINITION

AK049671 4435 bp mRNA linear HTC 05-DEC-2002
Mus musculus 12 days embryo spinal cord cDNA, RIKEN full-length
enriched library, clone:C530020122 product:GROWTH HORMONE
SECRETAGOGUE RECEPTOR TYPE 1 homolog [Rattus norvegicus], full
insert sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

Carninci, P., Shibata, Y., Hayata, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source

Location/Qualifiers

1. .608

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="A430103P18"

/tissue_type="thymus"

/dev_stage="0 day neonate"

/lab_host="DH10B"

/clone_lib="RIKEN full-length enriched, 0 day neonate thymus"

/notes="Site 1: Sali; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGGAGAGATTCGAGTAAATTAATATCCCCCCCCCCCC sequence [5' GAGGAGAGATTCGAGTAAATTAATATCCCCCCCCCCCC 3']. cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGGAGAGATTCGAGTAAATTAATATCCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

BASE COUNT

ORIGIN

Query Match 15.8%; Score 219.2; DB 14; Length 608;
Best Local Similarity 70.3%; Pred. No. 2e-35;
Matches 293; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

137 TGTGCTGCTTTCGTCGGGTGAGCGGCAAGTGTGACCGTGTATCGTGTATCGGGCGCT 196
2 TGGCGCTCTTCGTGGTGGGATCTCGGGCAACCTGCTCACCATGCTGGTGGTTCGGCT 61

197 ACCGGGACATCGGACACCAACAACTGTACCTGGGACAGTGGCGCTGTCGACCTAC 256
62 TCGCGAGCTGGGACACCAACCAACCTTACCTATCCATCATGGCCCTTCGATCTGC 121

257 TCATCTGCTCGGCTCGCGTTCGACTGTACCGCTCTTCGCGCTTCGCGCCCTCGGCTG 316
122 TCATCTCTCTGTCATCGCGCTGGACCTCTCGCGCTCTGGCAGTATCGGCCCTGGAAT 181

317 TCGGCGCGCTCTCGGCTGTCTCTACGTGTGCGGCGAGGCTGACCTAGCCACGC 376
182 TCGGCGACCTGCTCTGCAAACTCTTCAGTTGTTCAGCGAGAGCTGCACCTAGCCACGG 241

377 TGTGTCATGACGACCGGCTCAGCGTCGAGCGCTACCTGGGCACTTCGCGCCGCTCGCG 436
242 TCCTCACCATCAGCGCTGAGGCTGAGGCTACTTCGCCATCTGCTTCGCGCTGGGG 301
437 CCGCGCTTGTGTCACCGGCGCGGCTTCGCGCGCTCTATCGTGTGCTTCGGCGCGTGG 496

Db 302 CCAAGGTGGTGTTCACCAAGGGCCGCTGTGAAGCTGTGTCATCTTGTGTCGCGCGTGG 361
QY 497 CGCTGCTCTCTGCGCGTCCCTCTCTGCTGTGGCGTGCAGGAGGACCCCGCA 553
Db 362 CCTTCTGACGCGCGGCCCATCTTCNTGCTGTGGCGTGGAGACGACGACGCA 418

RESULT 4

BF513101

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgabs@mail.nih.gov

Oligo-dT track not found, Not 1 site shown in beginning of sequence

is likely internal to the message. cDNA library Preparation: M.B.

Soares lab Clone distribution: NCI-CGAP clone distribution

Information can be found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Seq primer: M13 Forward

POLYA=NO.

FEATURES

source

Location/Qualifiers

1. .590

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:3070549"

/lab_host="DH10B (Life Technologies)"

/clone_lib="NCI CGAP Sub7"

/notes="Vector: pT773D-Pac (Pharmacia) with a modified

polylinker; Site 1: Not I; Site 2: Eco RI; NCI CGAP Sub7

is a subtracted library derived from NCI CGAP Sub6. The

NCI CGAP Sub7 library had 12 million recombinants. A

single-stranded DNA preparation of NCI CGAP Sub6 was used

as a tracer in a subtractive hybridization with a driver

comprising: the IMAGE pool (NCI CGAP Kid3 pool 1 LLAM

3334-3337, 3682-3683, 3798-3803 (IMAGE Clones)

132376-132391, 145608-145677, 1500552-1502853);

NCI CGAP Kid5 pool 1 LLAM 3338-3342, 3722-3725, 3776-3778

(IMAGE Clones 1323912-1325831, 1471368-1472903,

1492104-1493255); NCI CGAP LUS pool 1 LLAM 3575-3582,

3851-3854 (IMAGE Clones 1414920-1417991, 1520904-1522439

); NCI CGAP GC4 pool 1 LLAM 3164-3167, 3716-3720,

3733-3735 (IMAGE Clones 1257096-1258631, 1469064-1470983

, 1475592-1476743); NCI CGAP Pr22 pool 1 LLAM 2457-2459,

2758-2759, 3062-3068 (IMAGE Clones 985608-986759,

1101192-1101959, 1217928-1220615); NCI CGAP Co10 pool 1

LLAM 2644-2653, 2871-2872 (IMAGE Clones 1057416-1061255

, 1144584-1145351). (6% of the driver population), plus a

pool of 3,840 arrayed clones from NCI CGAP Sub1 (IMAGE

Clones 2708616-2710535) and NCI CGAP Sub2 (IMAGE

Clones 2710536-2712455) (4% of the driver population

), plus a pool of 11,136 clones from NCI CGAP Sub3 (IMAGE

Clones 2712456-2723591) (10% of the driver population),

plus a pool of 5,472 clones from NCI CGAP Sub4 (IMAGE

Clones 2723592-2729326) (40% of the driver population),

plus a pool of 4032 clones from NCI CGAP Sub6 (IMAGE

Clones 2728969-2733190) (40% of the driver population).

Subtraction was performed as previously described [Bonaldo

, Lennon & Soares (1996): Normalization and Subtraction:

Two Approaches To Facilitate Gene Discovery. Genome


```

Research 6, 791-806.
TAG_SEQ=None found"
BASE COUNT      130 a      157 c      130 t
ORIGIN

Query Match      15.7% Score 218; DB 10; Length 590;
Best Local Similarity 100.0%; Pred. No. 3.6e-35;
Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 835 CGGCGCTGCGAGCGCGCGCGCTCGGGGGGAGAGAGCCACCGGAGAGCGTCCGC 894
DB 1 CGGCGCTGCGAGCGCGCGCGCTCGGGGGGAGAGAGCGCGGAGAGCGTCCGC 60
QY 895 GTCTGCGTGAAGTGAGCGCGCGTGTTCAGAGCGCTGCTGCAAGTCCGCGCGCG 954
DB 61 GTCTGCGTGAAGTGAGCGCGCGTGTTCAGAGCGCTGCTGCAAGTCCGCGCGCG 120
QY 955 GGGACCGCGCAAGCGTGGTCCCTTCCCTGCTGCTGCGCGCGCGCTTCCA 1014
DB 121 GGGACCGCGCAAGCGTGGTCCCTTCCCTGCTGCTGCGCGCGCGCTTCCA 180
QY 1015 GCTCCCTTCTCTATTTTCAGTTCCAGCTTCCACCGCGCG 1052
DB 181 GCTCCCTTCTCTATTTTCAGTTCCAGCTTCCACCGCGCG 218

RESULT 5
BU553576      843 bp mRNA linear EST 16-SEP-2002
LOCUS
DEFINITION AGENCOURT_10242213 NIH_MGC_109 Homo sapiens cDNA clone
IMAGE:6577973 5', mRNA sequence.
ACCESSION BU553576
VERSION BU553576.1 GI:22903848
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 843)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2778 row: p column: 05
High quality sequence stop: 534.
FEATURES
Location/Qualifiers
1..843
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6577973"
/tissue_type="teratocarcinoma, cell line"
/lab_hosts="DH10B (phage-resistant)"
/clone_lib="NIH MGC 109"
/notes="Organ: ovary; Vector: pOTB7; Site 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH MGC Library."
BASE COUNT      213 a      212 c      247 t      2 others
ORIGIN

Query Match      13.4% Score 186.4; DB 13; Length 843;
Best Local Similarity 91.0%; Pred. No. 1.5e-28;
Matches 243; Conservative 0; Mismatches 17; Indels 7; Gaps 4;

QY 1057 GGTGGTTCTGGCATTTTATAATTGCTGGTTGCCCTTCCACGTTGGCAGATCATTTACAT 1116
DB 534 GGTGGTTCTGGCATTTTATAATTGCTGGTTGCCCTTCCACGTTGGCAGATCATTTACAT 593
QY 1117 AAACACGGAAGATTCGCGGATGATGATCTCTCTCAGTACTTTTAACATCGTCGCTCGCA 1176
DB 594 AAACACGGAAGATTCGCGGATGATGATCTCTCTCAGTACTTTTAACATCGTCGCTCGCA 653
QY 1177 ACTTTTCTATCTGAGCGCATCTATCAACCAATCTCTCAACCTC-ATTTCAAAGAAGT 1235
DB 654 ACTTTTCTATCTGAGCGCATCTATCAACCAATCTCTCAACCTCAATTTTCAAGAAGT 713
QY 1236 AC-AGAGCGCGCGCTTTAAA---CTGCTGCTCGCAAGGAAGTCCAGGCCGAG-AGGCT 1389
DB 714 ACAAGAGCGCGCGCTTTAAAAGTCTGCTTGCAGAGGAAGTCCAGGCCGAGAGGCT 773
QY 1290 TCCACAGAGACGAGGACACTGCGGGGG 1316
DB 774 CCACAGAGACGAGGACACTGCGGGGG 800

RESULT 6
BU568940/c
LOCUS
DEFINITION AGENCOURT_10400272 NIH_MGC_82 Homo sapiens cDNA clone IMAGE:6616239
IMAGE:6577973 5', mRNA sequence.
ACCESSION BU568940
VERSION BU568940.1 GI:22919240
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 849)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: CLONTECH
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2859 row: j column: 15
High quality sequence stop: 429.
FEATURES
Location/Qualifiers
1..849
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6616239"
/lab_hosts="DH10B (T1 phage-resistant)"
/notes="Organ: testis; Vector: pDNR-LIB (Clontech); Site 1:
SfiI (ggccattatggcc); Site 2: SfiI (ggccattatggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGCCCATTTATGGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGGCGGCGGACATG-df(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size
1.35 kb (range 0.9-4.0 kb). 14/15 colonies contained
inserts by PCR. This library was enriched for full-length
clones and was constructed by Clontech Laboratories (Palo
Alto, CA)."
BASE COUNT      199 a      229 c      193 g      228 t
ORIGIN

```

Query Match 12.5%; Score 173.4; DB 13; Length 849;
 Best Local Similarity 88.6%; Pred. No. 7.3e-26;
 Matches 233; Conservative 0; Mismatches 25; Indels 5; Gaps 4;

QY 1132 GCGGATGATCTCTCTCAGTACTTAA-CATCGTCGCTCG--CAACTTTTCTATCT 1188
 DB 698 GGAAGAGGAGTACTTTCTCAGTACTTTAACTTGTGCGTTTGGCAACTTTTTTTTGT 639

QY 1189 GAGCGATCTATCAACCAATCTCTTACAACTTCAAAAGAGTACAGAG-CGCGCG 1247
 DB 638 AGGGCATTTATCAACCAATCTTACACTTCAATTTCAAGAGTACAGAGCGCGG 579

QY 1248 CTTTAACTGCTGCTCGAAGAGTCCAGGCGAGAGCTTCCACAGAGCAGGACA 1307
 DB 578 CTTTAACTGCTGCTCGAAGAGTCCAGG-CGAGAGGCTTCCACAGAGCAGGACA 520

QY 1308 CTGCGGGGAGTTCGAGGGGACACTGAGGAGACACGCTGGGCTACACGAGCAAGCG 1367
 DB 519 CTGCGGGGAGTTCGAGGGGACACTGAGGAGACACGCTGGGCTACACGAGCAAGCG 460

QY 1368 CTAACGTGAAGACGATGGGATAA 1390
 DB 459 CTAACGTGAAGACGATGGGATAA 437

RESULT 7
 AZ451922/c
 LOCUS 504 bp DNA linear GSS 04-OCT-2000
 DEFINITION IM0251C06R Mouse 10kb plasmid UUGCM library Mus musculus genomic
 clone UUGCM0251C06 R, genomic survey sequence.

ACCESSION AZ451922
 VERSION AZ451922.1 GI:10608187
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 504)
 Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
 Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
 M., Rose, M., Rose, R., Stokes, R., Fingey, A., von Niederhausen, A.,
 and Wright, D., Weiss, R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts

JOURNAL Unpublished
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0251 row: C column: 06
 Seq primer: CACACGAGAAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 504.
 Location/Qualifiers
 1..504
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGCM0251C06"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
 /clone_lib="Mouse 10kb plasmid UUGCM library"
 /notes="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 156 a 110 c 107 g 131 t
 ORIGIN

Query Match 10.2%; Score 141.2; DB 28; Length 504;
 Best Local Similarity 78.7%; Pred. No. 3.2e-19;
 Matches 181; Conservative 0; Mismatches 48; Indels 1; Gaps 1;

QY 1052 GTGGTGGTCTGCGCATTTAATTGC-TGGTTGCCCTCCAGTTGGCAGAAATCAT 1110
 DB 250 GTGGTGGAGGCTTGGCATTCATAGTTGCTGGTTGGCTTCCAAAGTTGGCAGAAATCT 191

QY 1111 TTACATAAACACGGAAGATTCCGGGATGATGTTCTCTCAGTACTTTAAACATCGTCGC 1170
 DB 190 TTTCATGAACACTGAAATTCGGGGTAAATGCACTCTCTCGATATCTTAATATGTCAC 131

QY 1171 TCTGCAACTTTTCTATCTGAGCGCATCTATCAACCAATCTCTTACACCTCATTTTCAA 1230
 DB 130 TCTCCAACTTCTATACCGAGTGCATCCATCAAGCCCATCTCCCAACCTCATGTCCAA 71

QY 1231 GAAGTACAGACGGCGGCTTTTAACTGCTGCGCAAGAGTCCAGGC 1280
 DB 70 TTAGTACAAGACGCTGTCCAGAACTGCTTTCGCCGAGACATGCCAGGC 21

RESULT 8
 BI757121
 LOCUS 855 bp mRNA linear EST 25-SEP-2001
 DEFINITION 603030882F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5200868 5',
 mRNA sequence.

ACCESSION BI757121
 VERSION BI757121.1 GI:15748699
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 855)
 NIH-MGC <http://mgi.nci.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 Plate: LLAM11503 row: d column: 21
 High quality sequence stop: 803.
 Location/Qualifiers
 1..855
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5200868"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_114"

FEATURES
 source


```

ACCESSION   BG714306
VERSION     BG714306.1  GI:13993237
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 634)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-r@mail.nih.gov
            cDNA Library Preparation: Miklos Palkovits, M.D., Ph.D.
            Tissue Procurement: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10671 row: g column: 03
            High quality sequence stop: 633.

FEATURES             source
    1..634
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:4792730"
        /tissue_type="hypothalamus"
        /lab_host="DH10B"
        /clone_lib="NIH MGC 96"
        /note="Organ: brain; Vector: pBluescriptR (modified
        pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag
        ); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3',
        size-selected for average insert size 2.3 kb and
        normalized to ROT 5. This is a primary library enriched
        for full-length clones and constructed using the
        Cap-trapper method (Carninci, in preparation). Library
        constructed by M. Brownstein (NIH/NHGRI, National
        Institutes of Health). Note: this is a NIH_MGC Library."
    74 a 221 c 243 g 96 t

BASE COUNT          74 a 221 c 243 g 96 t
ORIGIN
Query Match          9.0%; Score 125.4; DB 10; Length 634;
Best Local Similarity 55.0%; Pred. No. 6.4e-16;
Matches 293; Conservative 0; Mismatches 231; Indels 9; Gaps 2;

QY 27 CGGCCCCGAGGGGCGCGGAGCGCGCGTGGCCCGGCTGCGCGCTTGCAGCAGGCGCG 86
DB 49 CAGCAGCGCGCGCGCGCGCGCGCGCGCTCCAAACCGGGGCTGAGCTGGACCGCGGCT 108
QY 87 CTGCTGCGCTTTCCCTTGGGGCGCTGTGTGCGGTGACCGCTGTGTGCTGTGCTGTGTT 146
DB 109 GGGCGTGGACACTCGCTCTGGGCGAAGTGTCTTCCCGCGCTCTACGCACTCATCTG 168
QY 147 CGTCTGCGGGTGGCGGCAACGTG-----GTGACCGTATCTGATCGGGCGCTACCG 200
DB 169 GGGCGTGGGCGCGCGGGCGAATGCGTGTCCGTGCACGTGTGTGAAGCGCGGGCGCG 228
QY 201 GGACATCGGACACCAACCACTTGTACCTTGGGCGAGCATGGCGGTGTCGACCTACTCAT 260
DB 229 GCGAGCGGGCGGCTGCGGCGACACACGTGCTCAGCCTGGCGCTCGCGGGCTGTGTGCT 288
QY 261 CTGCTGCGGCTGCGGCTTGCACCTGTACCGCTTGTGCGCTCGCG---CCCTGGGTGTT 317
DB 289 GCTGTGCGGCTGCGGCTGAGCTCTACAGTTCGTGTGTTCCACTACCGCTGGTCTT 348
QY 318 CGGGCGGCTGCTGCGGCTGTCCCTCTACGTGGGCGAGGCTGACCTACGCGAGCT 377
DB 349 CGGCGACCTGGGTGCGCGGCTACTACTTTCGTGACGAGCTGTGCGGCTTACGCCACGGT 408
QY 378 GCTGCACATACCGCGGCTGACCGTTCGAGCGCTACCTGGCCATCTGCGCGCGCTCGCGC 437

```

```

DB 409 GCTGAGCGTGGCAGGCGCTGAGCGCGCGCTGCGCTAGCGCGTGTGCGAGCCCTGCGTGC 468
QY 438 CCGCGTCTTGTGTCACCGCGCGCGCGTCCGCGCGCTCATCGCTGTGCTCTGGGCGCGTGGC 497
DB 469 CCGCAGCCTCTGACCGCACCGCGGACCGCGTGGTGGCTCTGTTGGGCGCGCTC 528
QY 498 GCTGCTCTCTGCGCGTCCCTTCTTGTTCCTGTTGGGCGTGCAGCAGGACCCCG 550
DB 529 GCTCGGCGCTCGCCATGCGCCATGCGCGTCTCATGCGGCGAGACGACGAACCTCG 581

RESULT 11
BM0808829
LOCUS       BM0808829
DEFINITION AGENCOURT_6617918 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5734336
            5', mRNA sequence.
ACCESSION   BM0808829
VERSION     BM0808829.1  GI:19125652
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1055)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-r@mail.nih.gov
            Tissue Procurement: Invitrogen
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12739 row: h column: 17
            High quality sequence stop: 589.

FEATURES             source
    1..1055
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:5734336"
        /tissue_type="hippocampus"
        /lab_host="DH10B"
        /clone_lib="NIH MGC 124"
        /note="Organ: brain; Vector: pCMV-SPORT6; Site 1: EcoRV
        (destroyed); Site 2: NotI; RNA source male hippocampus,
        age 27. Library is oligo-dT primed and directionally
        cloned (EcoRV site is destroyed upon cloning). Average
        insert size 1.4 kb, insert size range 0.9-4 kb. Library is
        normalized and enriched for full-length clones and was
        constructed by C. Gruber (Invitrogen). Research Genetics
        tracking code 012."
    160 a 376 c 336 g 183 t

BASE COUNT          160 a 376 c 336 g 183 t
ORIGIN
Query Match          9.0%; Score 125.4; DB 12; Length 1055;
Best Local Similarity 55.0%; Pred. No. 7.3e-16;
Matches 293; Conservative 0; Mismatches 231; Indels 9; Gaps 2;

QY 27 CGGCCCCGAGGGGCGCGGAGCGCGCGTGGCCCGGCTGCGCGCTTGCAGCAGGCGCG 86
DB 53 CAGCAGCGCGCGCGCGCGCGCGCGCGCTCCAAACCGGGGCTGAGCTGGACCGCGGCT 112
QY 87 CTGCTGCGCTTTCCCTTGGGGCGCTGTGTGCGGTGACCGCTGTGCTGTGCTGTGTT 146
DB 113 GGGCGTGGACACTCGGCTCTGGGCGAAGTGTGTTTCCCGGCTCTACGCACTCATCTG 172
QY 147 CGTCTGCGGGTGGCGGCAACGTG-----GTGACCGTATCTGATCGGGCGCTACCG 200

```



```

source
1. .1027
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="NIH_MGC_124"
/notes="Organ: Brain; Vector: pCMV-SPORT6; Site 1: EcoRV
(deprecated); Site 2: NotI; RNA source male hippocampus,
age 27. Library is oligo-dT primed and directionally
cloned (EcoRV site is destroyed upon cloning). Average
insert size 1.4 kb, insert size range 0.9-4 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 012."
BASE COUNT 153 a 357 c 337 g 180 t
ORIGIN
Query Match 9.0%; Score 124.8; DB 12; Length 1027;
Best Local Similarity 56.0%; Pred. No. 9.6e-16;
Matches 281; Conservative 0; Mismatches 212; Indels 9; Gaps 2;

QY 58 CCGCGCTGCGGCTTGCAGAGCGCGCTGCTCGCCCTTCCCTGGGGCGCTGGTG 117
Db 83 CACCCGGGCTGAGCCTGGAGCGCCGCTGGCGCTGGACACTCGCCTTGGGCCAAGTG 142
QY 118 CCGGTGACCGCTGTGCGCTGTCCTGTTGCTGCTGGGTGAGCGGCAACGTG----- 171
Db 143 CTGTTACCGCGCTCTACGCACTCATCTGGGCGCTGGGCGCGGCAATGCGCTGTCC 202
QY 172 GTACACGTGATGTGATCGGGCGCTACCGGACATCGGACACCAACCACTTGTACCTG 231
Db 203 GTGACGTGCTGTAAGGCGCGGCGCGCGCGCGCGCTGGCGCCACCACTGTGCTC 262
QY 232 GGACGATGCGCGCTGTCGACCTACTCATCTGCTGGGCTGCGGTTCGACCTGTACCGC 291
Db 263 AGCTGGCGCTCGGGCGCTGCTGCTGCTGCTGGCGGTGCGGTGAGCTCTACAGC 322
QY 292 CTCTGGCGCTCGGG---CCCTGGGTGTTGGGCGCTGCTGCTGGCGCTGCTGCTTAC 348
Db 323 TTGCTGTGTTCCACTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTTC 382
QY 349 GTGGCGAGGCTGCACTACGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 408
Db 383 GTGACGAGCTGTGCGCTACGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 442
QY 409 TACCTGGCCATCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 468
Db 443 TGCTAGCGCTGTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 502
QY 469 GCGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 528
Db 503 TGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 562
QY 529 GTGGCGCTGACGAGGACCCCG 550
Db 563 ATGGGCGAGAAGCAAGCACTCG 584

RESULT 14
BM546496
LOCUS
DEFINITION
BM546496 1144 bp mRNA linear EST 20-FEB-2002
AGENCOURT_6494007 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5729257
5', mRNA sequence.
ACCESSION
BM546496
VERSION
BM546496.1 GI:18779503
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1144)

```

```

AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: Invitrogen
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLAM12726 row: e column: 02
High quality sequence start: 27
High quality sequence stop: 519.
FEATURES
Location/Qualifiers
1. .1144
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5729257"
/tissue_type="hippocampus"
/lab_host="DH10B"
/clone_lib="NIH_MGC_124"
/notes="Organ: Brain; Vector: pCMV-SPORT6; Site 1: EcoRV
(deprecated); Site 2: NotI; RNA source male hippocampus,
age 27. Library is oligo-dT primed and directionally
cloned (EcoRV site is destroyed upon cloning). Average
insert size 1.4 kb, insert size range 0.9-4 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 012."
BASE COUNT 188 a 384 c 371 g 200 t 1 others
ORIGIN
Query Match 8.9%; Score 124; DB 12; Length 1144;
Best Local Similarity 55.2%; Pred. No. 1.5e-15;
Matches 288; Conservative 0; Mismatches 225; Indels 9; Gaps 2;

QY 27 CCGCCCCGAGGGGCGCGGAGCGCGCTGGCGCGCTGGCGCGCTGGACGAGCGCG 86
Db 67 CAGCAGCGCGCGCGCGCTGGCGCGCTGGCGCGCTGGCGCGCTGGCGCGCGCGCT 126
QY 87 CTGCTCGCGCTTTCCTCGGGGCGCTGGCGCGCTGGCGCGCTGGCGCGCTGGCTGT 146
Db 127 GGGCGTGGACACTGCGCTCTGGGCGCAAGGTGCTTACCGCGCTCTACGACTCATCTG 186
QY 147 CGTCTGCGGGGTGAGCGCAACGTG-----GTGACCGTGTATGCTGATCGGGCGCTACCG 200
Db 187 GGGCGTGGGCGCGCGCGCGCAATGCGCTGCTCGGTGACAGTGGTGTGAAGCGCGCGCG 246
QY 201 GACATGCGGACCAACCACTTTGACCTGGGCGAGATGGCGGTGTCGACCTACTCAT 260
Db 247 GCGCGGGGCGCGCTGCGGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 306
QY 261 CTGCTCGGGCTGCGGCTTGCAGCTGTACCGCTC---TGGCGCTCGCGCGCTGGGT 317
Db 307 GCTGCTGCGGCTGCGGCTGAGCTTCTACAGCTTCTGTTGTTTCCATACCCCTGGGTCT 366
QY 318 CCGGCGCTGCTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 377
Db 367 CCGGCGCTGCTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 426
QY 378 GCTGCACTGACCGCGCTGAGCGCTGAGCGCTTACCTGGGCGATCTGGCGCGCGCTG 437
Db 427 GCTGAGCGTGGCAGGCGCTGAGCGCGCGGCGCTGCTGCTGCTGCTGCTGCTGCTG 486
QY 438 CCGCGCTTGGTCAACCGCGCGCGCTGCGCGCGCTCATGCTGCTGCTGCTGCTGCTG 497
Db 487 CCGCAGCTGCTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTC 546
QY 498 GCTGCTCTCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 539

```

Db 547 GCTCGGCTCGCCCTGCCCATGGCCGTATCATGCGGGGAGCA 588

RESULT 15
CNS036WH/c
LOCUS
DEFINITION
Tetraodon nigroviridis genome survey sequence T7 end of clone
216F22 of library G from Tetraodon nigroviridis, genomic survey
sequence.

ACCESSION
AL230570
VERSION
AL230570.1 GI:7889565
KEYWORDS
GSS; genome survey sequence.
SOURCE
Tetraodon nigroviridis
ORGANISM
Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
Tetraodontoidea; Tetraodontidae; Tetraodon.

REFERENCE
1 Roest Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizanes,C., Wincker,P., Brottier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
Estimate of human gene number provided by genome-wide analysis
using Tetraodon nigroviridis DNA sequence
Nat. Genet. 25 (2), 235-238 (2000)

REFERENCE
2 10835645

AUTHORS
Roest Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
Fizanes,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
Saurin,W., Bernot,A. and Weissenbach,J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
Genome Res. 10 (7), 939-949 (2000)

REFERENCE
3 20359837
10899143

MEDLINE
PUBMED

REFERENCE
3 (bases 1 to 941)
Genoscope.
Direct Submission
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/tetraodon.

FEATURES
source
Location/Qualifiers
1..941
/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="216F22"
/clone_lib="G"
/note="Genoscope sequence ID : COAG216DC11LP1-end : T7"

BASE COUNT
ORIGIN
197 a 277 c 327 g 132 t 8 others

Query Match
Best Local Similarity 8.8%; Score 122.8; DB 29; Length 941;
Matches 318; Conservative 2; Mismatches 270; Indels 5; Gaps 3;

Qy 24 CGACGGCCCCGAGGGGGCGCGGAGCGCCCGCGCGCTGCGCGCTTGCAGACGAGCG 83
Db 605 CGAGGCCACCGCGGGGACCTGGACGCGCTGTGACGAGGAGGCGTACTGGCCAG 546

Qy 84 CCGTGTGTGCGCCCTTTCCCTGGGGGGCGCTGGTCCCGGTGACCGCTGTGCTGTGCCT 143
Db 545 GTACCGGGGGCCCCCGAGTCCCGCTTCTCTGCGGTGCTGCGCTACCTGGCCAT 486

Qy 144 GTTCGTGTGCGGGGTGAGCGGCAACGTTGACCGTGTGCTGTGATCGGGCGCTACCGGGA 203
Db 485 CTTCTGTGGGGGTCTCTGGGCAACTCCCTGACCTGCGCGCTCATCTCTCGCTTCGCGGC 426

204 CATGCGGACCAACACCAACTTTGTACCTGGGACGATGGCGGTTCGACCTAC---TCAT 260
425 GAGCGCAGCGCCCAACCAACTACTACTCTGTGA-CCTGGCGCGCTCGACCTGTGTGCT 367
261 CTTGCTCGGGCTCGCGTTTGACCTGTGACCGCTCTGGCGCTCGCGGCGCTGGGTGTTGG 320
366 GCTGCTGGGGATGCCGCTGGAGCTGTACGACCTGTGGCGGAACCTACCCCTTCTGCTGG 307
321 GCCGCTGCTTGCCTGCTGCTTCCCTTACGTTGGCGGAGGGCTGCACCTACGCCACGCTGCT 380
306 GGAGGGGGGCTGCTACTTCCGGACCTTCTCTGTTTCGAG-ACGGTGTSTTTCGCTCCATCT 248
381 GCACATGACCGCGCTCAGCGTCCAGCGCTACTGCGGCATCTGCGCGCGCTCGCGCGCG 440
247 CAACGTCACGCGCTCAGCGTGGAGCGCTACCTGGCGGTGGTGCACCCGCTCAAGGTCAA 188
441 CGTCTTGTTCACCCCGCGCGCTCGCGCGCTCATCGCTGTGCTCTGGGCGCGTGGCGCT 500
187 GCACATGGGCACTCGCGCGCACGTCAAGAGGTGGTCTCTCTGCTGTGGCTGTGCTCAT 128
501 GCTCTCTGCGCGTCCCTTCTTCTTGTGTTGGCGCTCGAGACGACCCCGGCATCTCCGT 560
127 GCTGTGCGCGTGCCTCCCAACACACGAGCTGCACGCGATCGAGGSGTGGCGCCCAAGTTCGGC 68
561 AGTCCCGGGCTCAATGGCACCCCGCGGATCGCGCTCTCTCGCTCTCTCGCTCTCGTGG 615
67 CGGGAGTTTCTCTGCTCGCGCATCTGCTAGGTCCCTCGAACCTCTGTGCTCTGCTCG 13

Search completed: January 1, 2004, 02:52:42
Job time : 3261.52 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 1, 2004, 06:20:00 ; Search time 48.8546 Seconds
(without alignments)
1254.097 Million cell updates/sec

Title: US-09-719-485-5
Perfect score: 2043
Sequence: 1 MGSPWNGSDPGEPAPPP.....WQNLHKGKGFADVLLSVL 386

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2043	100.0	386	21	AAV54146
2	2043	100.0	386	22	AAV54146
3	1581	77.4	412	21	AAV54145
4	1581	77.4	412	22	AAV54145
5	1581	77.4	412	22	AAV54145
6	1581	77.4	412	23	AAV54145
7	1581	77.4	412	24	AAV54145
8	1581	77.4	501	22	AAV54145
9	1575	77.1	412	21	AAV54145

10	1573	77.0	412	23	ABB09535	Human motilin rece
11	1307	64.0	400	22	AAV54146	Amino acid sequenc
12	1149.5	56.3	271	22	AAV54146	Amino acid sequenc
13	673.5	33.0	363	21	AAV54147	The puffer fish mo
14	673.5	33.0	363	22	AAV54147	Amino acid sequenc
15	649	31.8	289	18	AAV19609	Pig growth hormone
16	649	31.8	289	18	AAV19611	Human growth hormo
17	649	31.8	289	18	AAV19612	Swine growth hormo
18	649	31.8	289	24	ABP81828	Human growth hormo
19	649	31.8	289	18	AAV19218	Human growth hormo
20	644	31.5	349	21	AAV69293	A canine growth ho
21	643	31.5	366	21	AAV90632	Human G protein-co
22	643	31.5	366	21	AAV70345	Rat growth hormone
23	643	31.5	366	22	AAV97376	Human G-protein co
24	643	31.5	366	22	AAV62650	Human ghrelin rece
25	643	31.5	366	23	ABB09534	A mouse growth hor
26	641	31.4	364	21	AAV54565	Rat growth hormone
27	641	31.4	364	22	AAV97377	Pig growth hormone
28	639.5	31.3	353	18	AAV19608	Swine growth hormo
29	639.5	31.3	353	18	AAV19215	Human growth hormo
30	639.5	31.3	361	18	AAV19217	Human growth hormo
31	639.5	31.3	362	18	AAV19610	Human growth hormo
32	637	31.2	364	18	AAV19613	Rat growth hormone
33	637	31.2	366	21	AAV90666	Human mutant G pro
34	625	30.6	364	18	AAV19220	Rat growth hormone
35	454	22.2	271	18	AAV19612	Human growth hormo
36	454	22.2	271	18	AAV19219	Human growth hormo
37	355.5	17.4	445	22	AAU68523	Human novel cytol
38	355.5	17.4	445	22	AAU68566	Human novel cytol
39	354.5	17.4	403	21	AAV90838	Human G protein-co
40	354.5	17.4	403	21	AAV44642	Human growth hormo
41	354.5	17.4	403	22	AAV99185	Human FM-3. Homo
42	354.5	17.4	403	23	ABG95158	Human GPCR GPR66
43	354.5	17.4	403	24	ABP81682	Human G-protein-co
44	354.5	17.4	426	22	AAE03628	Human G-protein co
45	352.5	17.3	403	21	AAV90673	Human mutant G pro

ALIGNMENTS

RESULT 1
AAV54146
ID AAV54146 standard; Protein; 386 AA.

XX AC AAV54146;

XX DT 27-MAR-2000 (first entry)

XX DE Amino acid sequence of the motilin receptor splice variant MTL-R1B.

XX KW Motilin receptor; G-protein coupled receptor; MTL-R1; GPR38;
KW spliced form; MTL-R1A; MTL-R1B; gastric motility disorder;
KW functional defect; neurological disorder; scleroderma; colonoscopy;
KW paraneoplastic syndrome; radiation induced dysmotility; diabetes;
KW infection; stress-related motility disorder; psychogenic disorder;
KW gastroparesis; gastro-oesophageal reflux disease; constipation;
KW chronic idiopathic pseudo obstruction; acute faecal impaction;
KW postoperative ileus; gallstones; infantile colic; diarrhoea;
KW irritable bowel syndrome; non-ulcer dyspepsia; non-cardiac chest pain;
KW endoscopy; duodenal intubation.

XX OS Homo sapiens.

XX PN WO964436-A1.

XX PD 16-DEC-1999.

XX PF 08-JUN-1999; 99WO-US12773.

XX PR 12-JUN-1998; 98US-0089098.

XX PA (MERI) MERCK & CO INC.

XX Feighner SD, Patchett AA, Tan C, McKee K, MacNeill D, Howard AD;
 PI Pong S, Smith RG;
 XX WPI; 2000-105868/09.
 DR N-PSDB; AAZ45404.
 XX Novel receptor protein for screening compounds used in treating
 PT irritable bowel syndrome, constipation and other gastric conditions -
 XX Claim 5; Fig 5; 44pp; English.
 XX The present sequence represents splice variant MTL-R1B of the motilin
 CC receptor. The gene encodes a G-protein coupled receptor, and is
 CC designated MTL-R1 (also GPR38). Two spliced forms of GPR38 exist,
 CC MTL-R1A (see AAY54145) and MTL-R1B (see AAY54146). MTL-R1A is a
 CC functional seven transmembrane domain form, and MTL-R1B is a truncated
 CC five transmembrane domain. The MTL-R1 proteins are used to identify
 CC agonists and antagonists which can be used for treating gastric motility
 CC disorders, functional defects, disorders secondary to neurological
 CC disorders e.g. scleroderma, paraneoplastic syndromes radiation induced
 CC dysmotility, diabetes, infections, stress-related motility disorders,
 CC psychogenic disorders, gastroparesis, gastro-oesophageal reflux disease,
 CC constipation, chronic idiopathic pseudo obstruction, acute faecal
 CC impaction, postoperative ileus, gallstones, infantile colic, irritable
 CC bowel syndrome, non-ulcer dyspepsia, non-cardiac chest pain and
 CC diarrhoea. They can also be used in the preparation for colonoscopy,
 CC endoscopy and duodenal intubation.
 XX Sequence 386 AA;
 SQ Query Match 100.0%; Score 2043; DB 21; Length 386;
 Best Local Similarity 100.0%; Pred. No. 2.9e-172;
 Matches 386; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGSPWNGSDGPEGAREPPWPPALPCDERRCSPPFLGALVPVTAVCLFLVVGSGNVTV 60
 Db 1 MGSPWNGSDGPEGAREPPWPPALPCDERRCSPPFLGALVPVTAVCLFLVVGSGNVTV 60
 QY 61 MLIGRYDRMTTNTNLYGSMVSDLLILGLPDLVRLWRSRPWFGPLLCRLSLVYVGE 120
 Db 61 MLIGRYDRMTTNTNLYGSMVSDLLILGLPDLVRLWRSRPWFGPLLCRLSLVYVGE 120
 QY 121 CTYATLLHMTALSVERYLAICRPLRVLVTRRRVRLIAVLWAVALLSAGPFLVGV 180
 Db 121 CTYATLLHMTALSVERYLAICRPLRVLVTRRRVRLIAVLWAVALLSAGPFLVGV 180
 QY 181 QDPGISVVPGLNGTARIASSPLASSPPLMLSRAPPPSPGPTAEAAALFSRECRPSA 240
 Db 181 QDPGISVVPGLNGTARIASSPLASSPPLMLSRAPPPSPGPTAEAAALFSRECRPSA 240
 QY 241 QLGLARVLMVWTTAYFFLPFLCISILYGLIGRELWSSRRPLRGPAAASGRGRHQT 300
 Db 241 QLGLARVLMVWTTAYFFLPFLCISILYGLIGRELWSSRRPLRGPAAASGRGRHQT 300
 QY 301 RKWSRRGSKDACLSQSPGTAQTGLPPLAQIWAPLPAPFPISIPASTRGGSGIYNL 360
 Db 301 RKWSRRGSKDACLSQSPGTAQTGLPPLAQIWAPLPAPFPISIPASTRGGSGIYNL 360
 QY 361 LVALPRQNLHLKHGREFADDVLLSVL 386
 Db 361 LVALPRQNLHLKHGREFADDVLLSVL 386
 RESULT 2
 AAB62653
 ID AAB62653 standard; Protein; 386 AA.
 XX
 AC AAB62653;
 XX
 DT 23-JUL-2001 (first entry)
 XX
 DE Short form of motilin receptor, GPR-38B isoform.

XX zsig33; signal transduction; hormone; enzyme; neural development;
 KW gastric contractility; nutrient uptake; digestive; pancreatic; human;
 KW insulin-like growth factor-I; growth hormone; bone; gastrointestinal;
 KW glucose; osteopathic; anorectic; vulnaray; immunomodulator; GHS-R;
 KW G-protein coupled receptor; motilin receptor; GPR38; GPR-38B; isoform.
 XX Homo sapiens.
 OS
 XX WO200138355-A2.
 PN
 XX 31-MAY-2001.
 PD
 XX 22-NOV-2000; 2000WO-US32074.
 PF
 XX 22-NOV-1999; 99US-0166765.
 PR
 XX (ZYMO) ZYMOGENETICS INC.
 PA
 XX Sheppard PO, Jaspers SR, Deisher TA, Bishop PD;
 PI WPI; 2001-355879/37.
 DR N-PSDB; AAF83684.
 XX Forming reversible peptide receptor complex for purifying cell and
 PT peptides, stimulating signal transduction and modulating hormone
 PT secretion, involves contacting a receptor with zsig33 polypeptide -
 XX Disclosure; Page 106-109; 11lpp; English.
 PS
 XX The invention relates to a method of forming a reversible peptide-
 CC receptor complex that involves providing an immobilized receptor, and
 CC contacting the receptor with a zsig33 peptide (comprising residues 24-37
 CC of AAB62649), where the receptor binds to the zsig33 peptide. The method
 CC is useful for purifying cells, purifying a peptide, stimulating signal
 CC transduction in a cell expressing a receptor. It is also useful for
 CC modulating secretion of hormones, neural development and/or utilization,
 CC gastric contractility, nutrient uptake, secretion of digestive and
 CC pancreatic enzymes and hormones, secretion of insulin-like growth factor
 CC -I, secretion of non-zsig33 proteins. It is useful for modulating growth
 CC hormone secretion in a mammal having a disease associated with abnormal
 CC levels of growth hormone, such as osteoporosis, bone repair, bone
 CC remodeling, low osteoblast levels, cartilage repair and remodeling,
 CC skeletal dysplasia, immune suppression, obesity, growth retardation,
 CC protein catabolic responses after surgery, cachexia, protein loss,
 CC dwarfism, wound healing and ovulation induction, treating a mammal having
 CC a metabolic disorder requiring neurological feedback, such as satiety
 CC regulation, glucose absorption and metabolism and neuropathy-associated
 CC gastrointestinal disorders, and stimulating glucose-induced insulin
 CC release in a mammal. The present sequence represents the short form of
 CC motilin receptor, GPR-38B (one of the two isoforms of GPR38 which result
 CC from alternative splicing). GPR38 has homology to the human G-protein
 CC coupled receptor, GHS-R.
 XX Sequence 386 AA;
 SQ Query Match 100.0%; Score 2043; DB 22; Length 386;
 Best Local Similarity 100.0%; Pred. No. 2.9e-172;
 Matches 386; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGSPWNGSDGPEGAREPPWPPALPCDERRCSPPFLGALVPVTAVCLFLVVGSGNVTV 60
 Db 1 MGSPWNGSDGPEGAREPPWPPALPCDERRCSPPFLGALVPVTAVCLFLVVGSGNVTV 60
 QY 61 MLIGRYDRMTTNTNLYGSMVSDLLILGLPDLVRLWRSRPWFGPLLCRLSLVYVGE 120
 Db 61 MLIGRYDRMTTNTNLYGSMVSDLLILGLPDLVRLWRSRPWFGPLLCRLSLVYVGE 120
 QY 121 CTYATLLHMTALSVERYLAICRPLRVLVTRRRVRLIAVLWAVALLSAGPFLVGV 180
 Db 121 CTYATLLHMTALSVERYLAICRPLRVLVTRRRVRLIAVLWAVALLSAGPFLVGV 180
 QY 181 QDPGISVVPGLNGTARIASSPLASSPPLMLSRAPPPSPGPTAEAAALFSRECRPSA 240

Db 181 QDPGISVVPGLNGTARTASSPLASSPPLWLSRAPPPSPGPGPETAFAALFSPRECRPSA 240
 Qy 241 QLGALRVMLVMTTAYFFLPCLCLSLYLGLIGRELWSSRRRLRGPAAAGRGHRQTVRVL 300
 Db 241 QLGALRVMLVMTTAYFFLPCLCLSLYLGLIGRELWSSRRRLRGPAAAGRGHRQTVRVL 300
 Qy 301 RKWSRRGSKDCLQSQAPPGTAQTGLPPLLAQWAPLPAPFPFISIPASTRRGGSGIYNL 360
 Db 301 RKWSRRGSKDCLQSQAPPGTAQTGLPPLLAQWAPLPAPFPFISIPASTRRGGSGIYNL 360
 Qy 361 LVALPRQNLHKGREADDVLLSVL 386
 Db 361 LVALPRQNLHKGREADDVLLSVL 386

RESULT 3

AAY54145

ID AAY54145 standard; Protein; 412 AA.

XX AC AAY54145;

XX DT 27-MAR-2000 (first entry)

XX DE Amino acid sequence of the motilin receptor splice variant MTL-R1A.

XX KW Motilin receptor; G-protein coupled receptor; MTL-R1; GPR38;
 KW spliced form; MTL-R1A; MTL-R1B; gastric motility disorder;
 KW functional defect; neurological disorder; scleroderma; colonoscopy;
 KW paraneoplastic syndrome; radiation induced dysmotility; diabetes;
 KW infection; stress-related motility disorder; psychogenic disorder;
 KW gastroparesis; gastro-oesophageal reflux disease; constipation;
 KW chronic idiopathic pseudo obstruction; acute faecal impaction;
 KW postoperative ileus; gallstones; infantile colic; diarrhoea;
 KW irritable bowel syndrome; non-ulcer dyspepsia; non-cardiac chest pain;
 KW endoscopy; duodenal intubation.

XX OS Homo sapiens.

XX PN WO9964436-A1.

XX PD 16-DEC-1999.

XX PF 08-JUN-1999; 99WO-US12773.

XX PR 12-JUN-1998; 98US-0089098.

XX PA (MERI) MERCK & CO INC.

XX PI Feighner SD, Patchett AA, Tan C, McKee K, MacNeil D, Howard AD;
 PI Pong S, Smith RG;

XX DR WPI; 2000-105868/09.

XX DR N-PSDB; AAZ45403.

XX PT Novel receptor protein for screening compounds used in treating
 PT irritable bowel syndrome, constipation and other gastric conditions

XX PS Claim 3; Fig 3; 44pp; English.

XX CC The present sequence represents splice variant MTL-R1A of the motilin
 CC receptor. The gene encodes a G-protein coupled receptor, and is
 CC designated MTL-R1 (also GPR38). Two spliced forms of GPR38 exist,
 CC MTL-R1A (see AAY54145) and MTL-R1B (see AAY54146). MTL-R1A is a
 CC functional seven transmembrane domain form, and MTL-R1B is a truncated
 CC five transmembrane domain. The MTL-R1 proteins are used to identify
 CC agonists and antagonists which can be used for treating gastric motility
 CC disorders, functional defects, disorders secondary to neurological
 CC disorders e.g. scleroderma, paraneoplastic syndromes radiation induced
 CC dysmotility, diabetes, infections, stress-related motility disorders,
 CC psychogenic disorders, gastroparesis, gastro-oesophageal reflux disease,
 CC constipation, chronic idiopathic pseudo obstruction, acute faecal
 CC impaction, postoperative ileus, gallstones, infantile colic, irritable

CC bowel syndrome, non-ulcer dyspepsia, non-cardiac chest pain and
 CC diarrhoea. They can also be used in the preparation for colonoscopy,
 CC endoscopy and duodenal intubation.

XX Sequence 412 AA;

XX Query Match 77.4%; Score 1501; DB 21; Length 412;

XX Best Local Similarity 100.0%; Pred. No. 2.2e-131;

XX Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGSPWNGSDGPGAREPPWALPPCDERRCSPPPLGALVPTAVCLCLFVGVSGNNVTV 60

Db 1 MGSPWNGSDGPGAREPPWALPPCDERRCSPPPLGALVPTAVCLCLFVGVSGNNVTV 60

Qy 61 MLIGRYRDMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWVFGPLLCRLSLVVGEG 120

Db 61 MLIGRYRDMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWVFGPLLCRLSLVVGEG 120

Qy 121 CTYATLLHMTALSVRYLAICRPLRARVLVTRRRVRLIAVLWAVALLSAGPPLFLVGVE 180

Db 121 CTYATLLHMTALSVRYLAICRPLRARVLVTRRRVRLIAVLWAVALLSAGPPLFLVGVE 180

Qy 181 QDPGISVVPGLNGTARTASSPLASSPPLWLSRAPPPSPGPGPETAFAALFSPRECRPSA 240

Db 181 QDPGISVVPGLNGTARTASSPLASSPPLWLSRAPPPSPGPGPETAFAALFSPRECRPSA 240

Qy 241 QLGALRVMLVMTTAYFFLPCLCLSLYLGLIGRELWSSRRRLRGPAAAGRGHRQTVRVL 300

Db 241 QLGALRVMLVMTTAYFFLPCLCLSLYLGLIGRELWSSRRRLRGPAAAGRGHRQTVRVL 300

RESULT 4

AAB62652

ID AAB62652 standard; Protein; 412 AA.

XX AC AAB62652;

XX DT 23-JUL-2001 (first entry)

XX DE Long form of motilin receptor, GPR-38A isoform.

XX KW zsig33; signal transduction; hormone; enzyme; neural development;
 KW gastric contractility; nutrient uptake; digestive; pancreatic; human;
 KW insulin-like growth factor-I; growth hormone; bone; gastrointestinal;
 KW glucose; osteopathic; anorectic; vulnery; immunomodulator; GHS-R;
 KW G-protein coupled receptor; motilin receptor; GPR38; GPR-38A; isoform.

XX OS Homo sapiens.

XX PN WO200138355-A2.

XX PD 31-MAY-2001.

XX PF 22-NOV-2000; 2000WO-US32074.

XX PR 22-NOV-1999; 99US-0166765.

XX PA (ZYMO) ZYMOGENETICS INC.

XX PI Sheppard PO, Jaspers SR, Deisher TA, Bishop PD;

XX DR WPI; 2001-355879/37.

XX DR N-PSDB; AAF83683.

XX PT Forming reversible peptide receptor complex for purifying cell and
 PT peptides, stimulating signal transduction and modulating hormone
 PT secretion, involves contacting a receptor with zsig33 polypeptide
 PS Disclosure; Page 102-104; 111pp; English.

XX CC The invention relates to a method of forming a reversible peptide-
 CC receptor complex that involves providing an immobilized receptor, and
 CC contacting the receptor with a zsig33 peptide (comprising residues 24-37

of AAB62649), where the receptor binds to the zsig33 peptide. The method is useful for purifying cells, purifying a peptide, stimulating signal transduction in a cell expressing a receptor. It is also useful for modulating secretion of hormones, neural development and/or utilization, gastric contractility, nutrient uptake, secretion of digestive and pancreatic enzymes and hormones, secretion of insulin-like growth factor -1, secretion of non-zsig33 proteins. It is useful for modulating growth hormone secretion in a mammal having a disease associated with abnormal levels of growth hormone, such as osteoporosis, bone repair, bone remodeling, low osteoblast levels, cartilage repair and remodeling, skeletal dysplasia, immune suppression, obesity, growth retardation, protein catabolic responses after surgery, cachexia, protein loss, dwarfism, wound healing and ovulation induction, treating a mammal having a metabolic disorder requiring neurological feedback, such as satiety regulation, glucose absorption and metabolism and neuropathy-associated gastrointestinal disorders, and stimulating glucose-induced insulin release in a mammal. The present sequence represents the long form of motilin receptor, GPR38A (one of the two isoforms of GPR38 which result from alternative splicing). GPR38 has homology to the human G-protein coupled receptor, GHS-R.

Query Match 77.4%; Score 1581; DB 22; Length 412;
Best Local Similarity 100.0%; Pred. No. 2.2e-131;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGSPWNGSDGEGAREPPWALPCDERRCSPPPLGALVPVTAVCILCFVVGSGNVTV 60
DB 1 MGSPWNGSDGEGAREPPWALPCDERRCSPPPLGALVPVTAVCILCFVVGSGNVTV 60
QY 61 MLIGRYDMRTTNNLYGSMVSDLLILGLPDLRLWRSRPPWFGPLLCRLSLYVGE 120
DB 61 MLIGRYDMRTTNNLYGSMVSDLLILGLPDLRLWRSRPPWFGPLLCRLSLYVGE 120
QY 121 CTYATLLHMTALSVERYLAICRPLRARVLTTRRRVRLIAVLWAVALLSAGPFLVGV 180
DB 121 CTYATLLHMTALSVERYLAICRPLRARVLTTRRRVRLIAVLWAVALLSAGPFLVGV 180
QY 181 QDGIISVVGNGTARTASSPLASSPPLWLSRAPPPSPGPGTAAALFSRECRSPA 240
DB 181 QDGIISVVGNGTARTASSPLASSPPLWLSRAPPPSPGPGTAAALFSRECRSPA 240
QY 241 QLGALRVMLVTTAYFFLPFLCISILYGLIGRLWSSRRPLRGPAASGRGRHQTQVRL 300
DB 241 QLGALRVMLVTTAYFFLPFLCISILYGLIGRLWSSRRPLRGPAASGRGRHQTQVRL 300

RESULT 5
AAB68478
ID AAB68478 standard; Protein; 412 AA.
XX AAB68478;
AC AAB68478;
DT 23-JUL-2001 (first entry)
XX Amino acid sequence of a human motilin receptor polypeptide.
DE Motilin receptor; gastrointestinal disease; gastric motility disorder;
KW gastroparesis; irritable bowel syndrome; diarrhoea.
XX Homo sapiens.
OS W0200132710-A1.
PN 10-MAY-2001.
XX 25-OCT-2000; 2000WO-US29426.
PF 29-OCT-1999; 99US-0162264.
XX (MERI) MERCK & CO INC.
XX

PI Tan C, McKee K;
XX WPI; 2001-343479/36.
DR N-PSDB; AAF85449.
DR Novel polypeptides related to dog and rabbit motilin receptor
XX polypeptide, comprising unique regions from dog and motilin receptor
PT amino acid sequence, useful for identifying compounds for treating
PT diarrhoea in humans
XX Disclosure; Page 32-33; 42pp; English.
XX The present sequence represents a human motilin receptor polypeptide.
XX The specification describes a unique sequence present in exon 1 of
CC the dog motilin receptor, which is not present in human or Sphaeroides
CC nephelus 7587 motilin receptor sequences. The unique nucleic acid
CC sequence is useful for measuring the ability of a compound to affect
CC motilin receptor activity. Motilin receptor polynucleotides and
CC polypeptides are used to identify therapeutic compounds which are
CC useful for treating gastrointestinal diseases and disorders such as
CC gastric motility disorders, gastroparesis, irritable bowel syndrome,
CC and diarrhoea.

Query Match 77.4%; Score 1581; DB 22; Length 412;
Best Local Similarity 100.0%; Pred. No. 2.2e-131;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGSPWNGSDGEGAREPPWALPCDERRCSPPPLGALVPVTAVCILCFVVGSGNVTV 60
DB 1 MGSPWNGSDGEGAREPPWALPCDERRCSPPPLGALVPVTAVCILCFVVGSGNVTV 60
QY 61 MLIGRYDMRTTNNLYGSMVSDLLILGLPDLRLWRSRPPWFGPLLCRLSLYVGE 120
DB 61 MLIGRYDMRTTNNLYGSMVSDLLILGLPDLRLWRSRPPWFGPLLCRLSLYVGE 120
QY 121 CTYATLLHMTALSVERYLAICRPLRARVLTTRRRVRLIAVLWAVALLSAGPFLVGV 180
DB 121 CTYATLLHMTALSVERYLAICRPLRARVLTTRRRVRLIAVLWAVALLSAGPFLVGV 180
QY 181 QDGIISVVGNGTARTASSPLASSPPLWLSRAPPPSPGPGTAAALFSRECRSPA 240
DB 181 QDGIISVVGNGTARTASSPLASSPPLWLSRAPPPSPGPGTAAALFSRECRSPA 240
QY 241 QLGALRVMLVTTAYFFLPFLCISILYGLIGRLWSSRRPLRGPAASGRGRHQTQVRL 300
DB 241 QLGALRVMLVTTAYFFLPFLCISILYGLIGRLWSSRRPLRGPAASGRGRHQTQVRL 300

RESULT 6
ABG30936
ID ABG30936 standard; Protein; 412 AA.
XX ABG30936;
AC ABG30936;
DT 21-OCT-2002 (first entry)
XX Human G protein-coupled receptor 38 (GPR38).
DE Human; G protein-coupled receptor 38; receptor; GPR38;
KW Alzheimer's disease; Parkinson's disease; ulcerative colitis;
KW Crohn's disease; Hodgkin's disease; glioblastoma; breast carcinoma;
KW colon carcinoma; lung small cell carcinoma; lung adenocarcinoma;
KW pancreatic small cell carcinoma; pancreatic adenocarcinoma.
XX Homo sapiens.
OS Key Location/Qualifiers
FH Region 15..31
FT /note= "Antigenic fragment"
FT 217..232
FT /note= "Antigenic fragment"

Region	276..291
FT	/note= "Antigenic fragment"
FT	373..388
Region	/note= "Antigenic fragment"
XX	
PN	WO200257791-A2.
XX	
PD	25-JUL-2002.
XX	
XX	29-NOV-2001; 2001WO-US45219.
XX	
PR	29-NOV-2000; 2000US-250251P.
PR	30-NOV-2000; 2000US-250452P.
XX	
PA	(LIFE-) LIFESPAN BIOSCIENCES INC.
XX	
XX	Brown JP, Burmer GC, Roush CL, Kulander BG;
DR	WPI; 2002-566812/60.
DR	N-FSDB; ABK90132.
XX	
PT	Assay for detecting Alzheimer's disease, Parkinson's disease,
PT	ulcerative colitis, Crohn's disease, Hodgkin's disease, glioblastoma or
PT	carcinoma, comprises using a binding partner for G protein coupled
PT	receptor 38 -
XX	
XX	Disclosure; Fig 2; 112pp; English.
XX	
CC	The present invention relates to a new assay method that involves
CC	contacting a binding partner specific for G protein coupled receptor
CC	(GPR) 38 with specific cells. The method of the invention is useful for
CC	the detection of an increased risk of Alzheimer's disease, Parkinson's
CC	disease, ulcerative colitis, Crohn's disease, Hodgkin's disease,
CC	glioblastoma, or carcinoma. GPR 38 is used to manufacture a medicament
CC	for inhibiting, treating or preventing Alzheimer's disease, Parkinson's
CC	disease, ulcerative colitis, Crohn's disease, Hodgkin's disease,
CC	glioblastoma, breast carcinoma, colon carcinoma, lung small cell
CC	carcinoma, lung adenocarcinoma, pancreatic small cell carcinoma, and
CC	pancreatic adenocarcinoma. An agonist or antagonist to GPR 38 are used
CC	to manufacture a medicament able to reduce the symptoms of these
CC	diseases. Nucleic acids encoding GPR 38 can also be used to treat the
CC	diseases. The present amino acid sequence represents the human G
CC	protein-coupled receptor 38 (GPR38) of the invention.
XX	
SQ	Sequence 412 AA;
Query Match	77.4%; Score 1581; DB 23; Length 412;
Best Local Similarity	100.0%; Prid. No. 2.2e-131;
Matches 300; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 MGSPWNGSDGEGAREPPWPALPCDERRCSPFPGALVPVTAVCLCFVVGVSGNVTV 60
Dd	1 MGSPWNGSDGEGAREPPWPALPCDERRCSPFPGALVPVTAVCLCFVVGVSGNVTV 60
Qy	61 MLIGRYDMRTTNLYLGSMVAVDLLILGLPFDLRLWRSRPWFGPLLCRLSLYVGEG 120
Dd	61 MLIGRYDMRTTNLYLGSMVAVDLLILGLPFDLRLWRSRPWFGPLLCRLSLYVGEG 120
Qy	121 CTYATILHMTALSVERYLAIICRPLARVLVTRRRVRALIAIWLVAALLSAGPFLVGVGE 180
Dd	121 CTYATILHMTALSVERYLAIICRPLARVLVTRRRVRALIAIWLVAALLSAGPFLVGVGE 180
Qy	181 QDPGISVVGINGTARTASSPLASSPPLWLGRAPPSPPSGPETAEAAALFSRECRPSA 240
Dd	181 QDPGISVVGINGTARTASSPLASSPPLWLGRAPPSPPSGPETAEAAALFSRECRPSA 240
Qy	241 QLGAIRVNLWTTAYFFLPFLCLLSLYGLIGRELWSSRRPLRGPAASGRERGHQTIVRVL 300
Dd	241 QLGAIRVNLWTTAYFFLPFLCLLSLYGLIGRELWSSRRPLRGPAASGRERGHQTIVRVL 300
RESULT 7	
ABP81993	

RESULT 7
ABP81993

Best Local Similarity 100.0%; Pred. No. 2.2e-131; Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		CC produce antibodies immunospecific for the polypeptides, and to identify membrane bound or soluble receptors.	
QY	1 MGSPWNGSDGEGAREPPWALPPCDERRCSPPPLGALVPVTAVCCLFVVGSGNVTV 60	SQ	Sequence 501 AA;
Db	1 MGSPWNGSDGEGAREPPWALPPCDERRCSPPPLGALVPVTAVCCLFVVGSGNVTV 60	Query Match	77.4%; Score 1581; DB 22; Length 501;
QY	61 MLIGRYDMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWVFGPLLCRLSLVYVGE 120	Best Local Similarity	100.0%; Pred. No. 2.7e-131;
Db	61 MLIGRYDMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWVFGPLLCRLSLVYVGE 120	Matches	300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	121 CTYATLLHMTALSVERYLAICRPLARVLVTRRRVRALIAVLWAVALLSAGPFLVGV 180	QY	1 MGSPWNGSDGEGAREPPWALPPCDERRCSPPPLGALVPVTAVCCLFVVGSGNVTV 60
Db	121 CTYATLLHMTALSVERYLAICRPLARVLVTRRRVRALIAVLWAVALLSAGPFLVGV 180	Db	90 MGSPWNGSDGEGAREPPWALPPCDERRCSPPPLGALVPVTAVCCLFVVGSGNVTV 149
QY	181 QDPGISVVGNGTARTASSPLASSPPLWLSRAPPPSPGPTAEAAALFSRECRPSA 240	QY	61 MLIGRYDMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWVFGPLLCRLSLVYVGE 120
Db	181 QDPGISVVGNGTARTASSPLASSPPLWLSRAPPPSPGPTAEAAALFSRECRPSA 240	Db	150 MLIGRYDMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWVFGPLLCRLSLVYVGE 209
QY	241 QLGALRVMLVMTTAYFFLPFLCISILYGLIGRELWSSRRPLRGPAASGRERGHRTVRL 300	QY	121 CTYATLLHMTALSVERYLAICRPLARVLVTRRRVRALIAVLWAVALLSAGPFLVGV 180
Db	241 QLGALRVMLVMTTAYFFLPFLCISILYGLIGRELWSSRRPLRGPAASGRERGHRTVRL 300	Db	210 CTYATLLHMTALSVERYLAICRPLARVLVTRRRVRALIAVLWAVALLSAGPFLVGV 269
RESULT 8		QY	181 QDPGISVVGNGTARTASSPLASSPPLWLSRAPPPSPGPTAEAAALFSRECRPSA 240
AAAG65822	ID AAG65822 standard; Protein; 501 AA.	Db	270 QDPGISVVGNGTARTASSPLASSPPLWLSRAPPPSPGPTAEAAALFSRECRPSA 329
AC	AAG65822;	QY	241 QLGALRVMLVMTTAYFFLPFLCISILYGLIGRELWSSRRPLRGPAASGRERGHRTVRL 300
XX	30-JAN-2002 (first entry)	Db	330 QLGALRVMLVMTTAYFFLPFLCISILYGLIGRELWSSRRPLRGPAASGRERGHRTVRL 399
DT	Human GPR38 variant GPR38V polypeptide.	RESULT 9	
DE	GPR38V; variant; antibacterial; cytostatic; analgesic; antiasthmatic;	AAAB02854	ID AAB02854 standard; Protein; 412 AA.
KW	anti-Parkinsonian; hypertensive; hypotensive; antidiabetic; osteopathic;	XX	AC AAB02854;
KW	antiallergic; antimigraine; neuroleptic; nootropic; anticonvulsant;	XX	DT 22-AUG-2000 (first entry)
KW	antiulcer; antiemetic; cardiant; vaccine; human.	XX	Human G protein coupled receptor hGPR38 (V297K) protein SEQ ID NO:130.
XX	Homo sapiens.	XX	Human; G protein coupled receptor; GPCR; transmembrane receptor;
XX	WO200164836-A2.	XX	identification; agonist; screening; therapeutic; pharmaceutical;
XX	07-SEP-2001.	XX	mutant.
XX	28-FEB-2001; 2001WO-US06277.	OS	Homo sapiens.
XX	01-MAR-2000; 2000US-0516315.	OS	Synthetic.
PA	(SMIK) SMITHKLINE BEECHAM CORP.	XX	WO200022131-A2.
PI	Elshourbagy N, Shabon U;	XX	20-APR-2000.
XX	WPI; 2001-638956/73.	XX	13-OCT-1999; 99WO-US24065.
XX	N-PSDB; AAI66989.	XX	13-OCT-1998; 98US-0170496.
XX	New human GPR38V polypeptide and polynucleotide, useful for treating	PR	12-NOV-1998; 98US-0108029.
PT	e.g. bacterial, fungal, protozoal and viral infections, cancers or	PR	20-NOV-1998; 98US-0109213.
PT	allergies, as vaccines, and for identifying agonists and antagonists	PR	27-NOV-1998; 98US-0110060.
PT	potentially useful in therapy	PR	16-FEB-1999; 99US-0120416.
XX	Claim 1; Page 26; 32pp; English.	PR	26-FEB-1999; 99US-0121852.
XX	This represents a human GPR38 variant (GPR38V) polypeptide. GPR38V can	PR	12-MAR-1999; 99US-0123944.
CC	be expressed by standard recombinant methodology. The polynucleotides and	PR	12-MAR-1999; 99US-0123946.
CC	polypeptides are used in the treatment of bacterial, fungal, protozoal	PR	12-MAR-1999; 99US-0123949.
CC	and viral infections, infections caused by HIV-1 or HIV-2, pain, cancers,	PR	12-MAR-1999; 99US-0123951.
CC	diabetes, obesity, anorexia, asthma, Parkinson's disease, acute heart	PR	28-MAY-1999; 99US-0136436.
CC	failure, hypertension, urinary retentions, osteoporosis, allergies,	PR	28-MAY-1999; 99US-0136437.
CC	ulcers, migraine, psychotic and neurological disorders, or dyskinesias.	PR	28-MAY-1999; 99US-0137127.
CC	They are also useful for identifying agonists and antagonists that are	PR	28-MAY-1999; 99US-0137131.
CC	potentially useful in therapy, as vaccines to induce immunological	PR	28-MAY-1999; 99US-0137567.
CC	response in a mammal. The polypeptides may also be used as immunogens to	PR	30-JUN-1999; 99US-0141448.
CC		PR	27-AUG-1999; 99US-0151114.

PR 03-SEP-1999; 99US-0152524.
 PR 29-SEP-1999; 99US-0156633.
 PR 29-SEP-1999; 99US-0156555.
 PR 29-SEP-1999; 99US-0156634.
 PA (AREN-) ARENA PHARM INC.
 XX Behan DP, Lehmann-Bruinema K, Chalmers DT, Chen R, Dang HT;
 PI Gore M, Liaw CW, Lin I, Lowitz K, White C;
 XX WPI: 2000-317986/27.
 DR N-PSDB; AAA46116.
 XX Non-endogenous, human G protein-coupled receptors for screening
 PT receptor, inverse or partial agonists useful as therapeutic agents -
 XX Example 2; Page 168-169; 187pp; English.
 XX The present invention describes transmembrane receptors, preferably
 CC human G protein coupled receptors (GPCR), for which the endogenous
 CC ligand is unknown (Orphan GPCR receptors). More specifically the present
 CC invention relates to non-endogenous, constitutively activated versions
 CC of a human GPCR. These non-endogenous human GPCRs can be useful for
 CC the direct identification of candidate compounds as receptors agonists,
 CC inverse agonists or partial agonists for use as pharmaceutical agents.
 CC AAA46017 to AAA46126 and AAA02825 to AAA02859 represent sequences used in
 CC the exemplification of the present invention.
 XX Sequence 412 AA;
 SQ
 Query Match 77.1%; Score 1575; DB 21; Length 412;
 Best Local Similarity 99.7%; Pred. No. 7.4e-131;
 Matches 299; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MGSPWNGSDGPEGAREPPWPPALPPCDERRCSPPFPGALVPVTVAVCLCLFVGVSGNVTV 60
 DB 1 MGSPWNGSDGPEGAREPPWPPALPPCDERRCSPPFPGALVPVTVAVCLCLFVGVSGNVTV 60
 QY 61 MLIGRYDMRTTNNLYLGSMVSDLLILLGLPFDLYRLWRSRPWVFGPPLCLSLYVGE 120
 DB 61 MLIGRYDMRTTNNLYLGSMVSDLLILLGLPFDLYRLWRSRPWVFGPPLCLSLYVGE 120
 QY 121 CYATLLHMTALSVERYLAICRPLRARVLVTRRRVRLIAVLMAVALLSAGPFLVGYE 180
 DB 121 CYATLLHMTALSVERYLAICRPLRARVLVTRRRVRLIAVLMAVALLSAGPFLVGYE 180
 QY 181 QDPGISVVPGLNGTARIASSPLASPPPLWLRAPPPSPGPTAEAAALFSRECRPSA 240
 DB 181 QDPGISVVPGLNGTARIASSPLASPPPLWLRAPPPSPGPTAEAAALFSRECRPSA 240
 QY 241 QLGALRVMLWVTYAFFFLPCLSLYGLIGRELWSSRRPLGPAASGRGRHROTIVRL 300
 DB 241 QLGALRVMLWVTYAFFFLPCLSLYGLIGRELWSSRRPLGPAASGRGRHROTIVRL 300
 RESULT 10
 ID ABB09535 standard; protein; 412 AA.
 XX ABB09535;
 AC ABB09535;
 XX 22-OCT-2002 (first entry)
 DE Human motilin receptor.
 XX Human; motilin receptor; appetite; food intake; agonist; analogue;
 KW undernutrition; anorexia; cachexia; malignant disease; infection;
 KW inflammatory disease; weight loss; antagonist; obesity; anorectic;
 KW anabolic; ghrelin receptor homologue; receptor.
 OS Homo sapiens.
 XX WO200260472-A1.

XX 08-AUG-2002.
 PD 31-JAN-2002; 2002WO-JP00765.
 PF 31-JAN-2001; 2001JP-0024423.
 PR (CHUS) CHUGAI SEIYAKU KK.
 XX Inui A, Asakawa A, Kaga T;
 PI WPI: 2002-619206/66.
 DR Remedies for diseases with hypo-nutrition status e.g. inappetence and
 PT cachexia, containing ghrelin or its analogs including agonists and
 PT antagonists -
 XX Disclosure; Fig 1B; 50pp; Japanese.
 XX The invention relates to the use of ghrelin or its analogues for the
 CC treatment of diseases associated with undernutrition such as anorexia,
 CC and also relates to the use of ghrelin antagonists for the prevention
 CC or treatment of obesity. The invention additionally discloses a method
 CC for screening ghrelin agonists or antagonists by measuring the amount
 CC of food intake, neuropeptide Y (NPY) expression, binding of NPY to
 CC NPY receptor Y1, oxygen consumption, gastric emptying, or activity of
 CC the vagus nerve. Intracerebroventricular (ICV) administration of ghrelin
 CC in animals was found to increase food intake over a period of 24 hours.
 CC Ghrelin and its analogues may therefore be used to treat conditions
 CC such as loss of appetite, anorexia, cachexia, malignant diseases, and
 CC weight loss associated with infection or inflammatory diseases.
 CC Conversely, ghrelin antagonists may be used in the treatment of obesity.
 CC The present sequence represents the human motilin receptor, a homologue
 CC of the ghrelin receptor (ABB09534) which is referred to in the disclosure
 CC of the invention.
 XX Sequence 412 AA;
 SQ
 Query Match 77.0%; Score 1573; DB 23; Length 412;
 Best Local Similarity 99.7%; Pred. No. 1.1e-130;
 Matches 299; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MGSPWNGSDGPEGAREPPWPPALPPCDERRCSPPFPGALVPVTVAVCLCLFVGVSGNVTV 60
 DB 1 MGSPWNGSDGPEGAREPPWPPALPPCDERRCSPPFPGALVPVTVAVCLCLFVGVSGNVTV 60
 QY 61 MLIGRYDMRTTNNLYLGSMVSDLLILLGLPFDLYRLWRSRPWVFGPPLCLSLYVGE 120
 DB 61 MLIGRYDMRTTNNLYLGSMVSDLLILLGLPFDLYRLWRSRPWVFGPPLCLSLYVGE 120
 QY 121 CYATLLHMTALSVERYLAICRPLRARVLVTRRRVRLIAVLMAVALLSAGPFLVGYE 180
 DB 121 CYATLLHMTALSVERYLAICRPLRARVLVTRRRVRLIAVLMAVALLSAGPFLVGYE 180
 QY 181 QDPGISVVPGLNGTARIASSPLASPPPLWLRAPPPSPGPTAEAAALFSRECRPSA 240
 DB 181 QDPGISVVPGLNGTARIASSPLASPPPLWLRAPPPSPGPTAEAAALFSRECRPSA 240
 QY 241 QLGALRVMLWVTYAFFFLPCLSLYGLIGRELWSSRRPLGPAASGRGRHROTIVRL 300
 DB 241 QLGALRVMLWVTYAFFFLPCLSLYGLIGRELWSSRRPLGPAASGRGRHROTIVRL 300
 RESULT 11
 ID AAB68477 standard; Protein; 400 AA.
 XX AAB68477;
 AC AAB68477;
 XX 23-JUL-2001 (first entry)
 DE Amino acid sequence of a rabbit motilin receptor polypeptide.
 XX

KW Motilin receptor; gastrointestinal disease; gastric motility disorder;
 KW gastroparesis; irritable bowel syndrome; diarrhoea.
 OS Oryctolagus cuniculus.
 XX WO200132710-A1.
 PN 10-MAY-2001.
 XX 25-OCT-2000; 2000WO-US29426.
 PF 29-OCT-1999; 99US-0162264.
 XX (MERI) MERCK & CO INC.
 PA Tan C, McKee K;
 PI WPI; 2001-343479/36.
 DR N-PSDB; AAF85448.
 XX Novel polypeptides related to dog and rabbit motilin receptor
 PT polypeptide, comprising unique regions from dog and motilin receptor
 PT amino acid sequence, useful for identifying compounds for treating
 PT diarrhoea in humans
 XX Claim 8; Page 18; 42pp; English.
 PS The present sequence represents a rabbit motilin receptor polypeptide.
 CC The specification describes a unique sequence present in exon 1 of
 CC the dog motilin receptor, which is not present in human or Sphaeroides
 CC nephelus 75E7 motilin receptor sequences. The unique nucleic acid
 CC sequence is useful for measuring the ability of a compound to affect
 CC motilin receptor activity. Motilin receptor polynucleotides and
 CC polypeptides are used to identify therapeutic compounds which are
 CC useful for treating gastrointestinal diseases and disorders such as
 CC gastric motility disorders, gastroparesis, irritable bowel syndrome,
 CC and diarrhoea.
 XX Sequence 400 AA;
 SQ Query Match 64.0%; Score 1307; DB 22; Length 400;
 Best Local Similarity 85.3%; Pred. No. 3.5e-107;
 Matches 256; Conservative 5; Mismatches 27; Indels 12; Gaps 3;
 QY 1 MGSPWNGSDGPGAREPPWPPALPCDERRCSPFPLGALVPVTAVCLCLFVVGSGNVTV 60
 Db 1 MGSPWNGSDGPGAREPPWPPALPCDERRCSPFPLGALVPVTAVCLCLFVVGSGNVTV 60
 QY 61 MLIGRYDMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWVFGPPLCLRLSLYVGE 120
 Db 61 LLIGRYDMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWVFGPPLCLRLSLYVGE 120
 QY 121 CTYATLLHMTALSVERYLAICRPLARLVTRRRVRLIAVLWAVALLSAGPFLFLVGVE 180
 Db 121 CTYASLLHMTALSVERYLAICRPLARLVTRRRVRLIAVLWAVALLSAGPFLFLVGVE 180
 QY 181 QDPGISVVGNGTARIASSPLASPPWLSRAPPSPPSGPETAAALFSRECRPSGA 240
 Db 181 QDPAPVAPDRNGTV-----PLDPSP-----APASPPSGP-GAAALFSRECRPSGA 228
 QY 241 QLGLRVMLWVTYAFFFLPFLCLSLYGLIGRELWSSRRPLRGPAAASGRGRHQTVRVL 300
 Db 229 QLGLRVMLWVTYAFFFLPFLCLSLYGLIAQLWRGRLGPAATGRGRHQTVRVL 288
 RESULT 12
 AAB68476
 ID AAB68476 standard; Protein; 271 AA.
 XX AAB68476;
 AC AAB68476;
 XX 23-JUL-2001 (first entry)
 DT XX

DE Amino acid sequence of a dog motilin receptor exon 1.
 XX Motilin receptor; gastrointestinal disease; gastric motility disorder;
 KW gastroparesis; irritable bowel syndrome; diarrhoea.
 OS Canis sp.
 XX WO200132710-A1.
 PN 10-MAY-2001.
 XX 25-OCT-2000; 2000WO-US29426.
 PF 29-OCT-1999; 99US-0162264.
 XX (MERI) MERCK & CO INC.
 PA Tan C, McKee K;
 PI WPI; 2001-343479/36.
 DR N-PSDB; AAF85447.
 XX Novel polypeptides related to dog and rabbit motilin receptor
 PT polypeptide, comprising unique regions from dog and motilin receptor
 PT amino acid sequence, useful for identifying compounds for treating
 PT diarrhoea in humans
 XX Claim 4; Page 17; 42pp; English.
 PS The present sequence is encoded by exon 1 of a dog motilin receptor gene.
 CC The specification describes a unique sequence present in exon 1 of
 CC the motilin receptor, which is not present in human or Sphaeroides
 CC nephelus 75E7 motilin receptor sequences. The unique nucleic acid
 CC sequence is useful for measuring the ability of a compound to affect
 CC motilin receptor activity. Motilin receptor polynucleotides and
 CC polypeptides are used to identify therapeutic compounds which are
 CC useful for treating gastrointestinal diseases and disorders such as
 CC gastric motility disorders, gastroparesis, irritable bowel syndrome,
 CC and diarrhoea.
 XX Sequence 271 AA;
 SQ Query Match 56.3%; Score 1149.5; DB 22; Length 271;
 Best Local Similarity 77.0%; Pred. No. 1.9e-93;
 Matches 231; Conservative 13; Mismatches 27; Indels 29; Gaps 4;
 QY 1 MGSPWNGSDGPGAREPPWPPALPCDERRCSPFPLGALVPVTAVCLCLFVVGSGNVTV 60
 Db 1 MGSPWNGSDGPGAREPPWPPALPCDERRCSPFPLGALVPVTAVCLCLFVVGSGNVTV 53
 QY 61 MLIGRYDMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWVFGPPLCLRLSLYVGE 120
 Db 54 LLIGRYDMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWVFGPPLCLRLSLYVGE 113
 QY 121 CTYATLLHMTALSVERYLAICRPLARLVTRRRVRLIAVLWAVALLSAGPFLFLVGVE 180
 Db 114 CTYATLLHMTALSVERYLAICRPLARLVTRRRVRLIAVLWAVALLSAGPFLFLVGVE 173
 QY 181 QDPGISVVGNGTARIASSPLASPPWLSRAPPSPPSGPETAAALFSRECRPSGA 240
 Db 174 QDAG---GFLNGSARLARA-----PSPPPGPE-----AALFSRECRPSPS 211
 QY 241 QLGLRVMLWVTYAFFFLPFLCLSLYGLIGRELWSSRRPLRGPAAASGRGRHQTVRVL 300
 Db 212 QLGLRVMLWVTYAFFFLPFLCLSLYGLIGRELWSSRRPLRGPAAASGRGRHQTVRVL 271
 RESULT 13
 AAY54147
 ID AAY54147 standard; Protein; 363 AA.
 XX AAY54147;
 AC AAY54147;
 XX

DT 27-MAR-2000 (first entry)
XX The puffer fish motilin receptor protein encoded by clone 75E7.
DE
KW Motilin receptor; G-protein coupled receptor; MTL-R1; GPR38;
KW spliced form; MTL-R1A; MTL-R1B; gastric motility disorder;
KW functional defect; neurological disorder; scleroderma; colonoscopy;
KW paraneoplastic syndrome; radiation induced dysmotility; diabetes;
KW infection; stress-related motility disorder; psychogenic disorder;
KW gastroparesis; gastro-oesophageal reflux disease; constipation;
KW chronic idiopathic pseudo obstruction; acute faecal impaction;
KW postoperative ileus; gallstones; infantile colic; diarrhoea;
KW irritable bowel syndrome; non-ulcer dyspepsia; non-cardiac chest pain;
KW endoscopy; duodenal intubation.
OS Sphaeroides nephelus.
XX WO9964436-A1.
PN 16-DEC-1999.
XX 08-JUN-1999; 99WO-US12773.
XX 12-JUN-1998; 98US-0089098.
XX (MERI) MERCK & CO INC.
PA Feighner SD, Patchett AA, Tan C, McKee K, MacNeil D, Howard AD;
PI Pong S, Smith RG;
PI Pong S, Smith RG;
XX WPI; 2000-105868/09.
DR N-PSDB; AA245405.
XX Novel receptor protein for screening compounds used in treating
PT irritable bowel syndrome, constipation and other gastric conditions -
PT Claim 7; Fig 8; 44pp; English.
PS The present sequence represents a motilin receptor (clone 75E7) which
CC is related to the human motilin receptor of the invention. The
CC specification describes a G-protein coupled receptor, designated
CC MTL-R1 (also GPR38). Two spliced forms of GPR38 exist, MTL-R1A (see
CC AAY54145) and MTL-R1B (see AAY54146). MTL-R1A is a functional seven
CC transmembrane domain form, and MTL-R1B is a truncated five
CC transmembrane domain. The MTL-R1 proteins are used to identify agonists
CC and antagonists which can be used for treating gastric motility
CC disorders, functional defects, disorders secondary to neurological
CC disorders e.g. scleroderma, paraneoplastic syndromes radiation induced
CC dysmotility, diabetes, infections, stress-related motility disorders,
CC psychogenic disorders, gastroparesis, gastro-oesophageal reflux disease,
CC constipation, chronic idiopathic pseudo obstruction, acute faecal
CC impaction, postoperative ileus, gallstones, infantile colic, irritable
CC bowel syndrome, non-ulcer dyspepsia, non-cardiac chest pain and
CC diarrhoea. They can also be used in the preparation for colonoscopy,
CC endoscopy and duodenal intubation.
XX Sequence 363 AA;
SQ Query Match 33.0%; Score 673.5; DB 21; Length 363;
Best Local Similarity 49.6%; Pred. No. 3.2e-51;
Matches 135; Conservative 39; Mismatches 55; Indels 43; Gaps 4;

QY 31 SPFPGALVPVAVCLFLVGVGNGVTVMLIGRYDMRTTNNLYLGSMAVSDDLILLG 90
DB 31 SLFPASTLPVTVICILIFVVGVTGMTLLIQFKDKMTTNLYLSMAVSDDLIVPLC 90

QY 91 LPFDLYRLWRSRPWFVGPGLLCRLSLYVGGCTGYATLLHMTALSVERYLAICRPLRARVLV 150
DB 91 LPFDLYRLWKYPWLPFGVAECRLHYIPGCTSAITLHTALSISERYLAISFPLRSKVMV 150

QY 151 TRRRVALLAVLWALLSAGPFLFLVGVGEQDGISSVVVGLNGTARIASSPLASSPPULWL 210
DB 151 TRRRVQIITLAWCFALVGAAPTFLVGVGEYO-----NET----- 185

QY 211 SRAPPPSPSPSETAAALFSRECRSPAOL--GALRVMLWVTATTATFFFLCLSLSYLG 268
DB 186 ----HPDYNTG-----QCKHTGYAISSQLHIMIWTSTVTFPCPMCLLFYLG 229

QY 269 LIGRELWSRRRLRGPAASGRBGHRQTVRVL 300
DB 230 SIGCKLMKSINDLOQCICALARESHROTVKIL 261

RESULT 14
AAB68479
ID AAB68479 standard; Protein; 363 AA.
XX AAB68479;
XX 23-JUL-2001 (first entry)
XX Amino acid sequence of a motilin receptor polypeptide.
XX Motilin receptor; gastrointestinal disease; gastric motility disorder;
KW gastroparesis; irritable bowel syndrome; diarrhoea.
XX Sphaeroides nephelus.
XX WO200132710-A1.
XX 10-MAY-2001.
XX 25-OCT-2000; 2000WO-US29426.
XX 29-OCT-1999; 99US-0162264.
XX (MERI) MERCK & CO INC.
XX Tan C, McKee K;
XX WPI; 2001-343479/36.
DR N-PSDB; AAF85450.
XX Novel polypeptides related to dog and rabbit motilin receptor
PT polypeptide, comprising unique regions from dog and motilin receptor
PT amino acid sequence, useful for identifying compounds for treating
PT diarrhoea in humans
XX Disclosure; Page 33-34; 42pp; English.
XX The present sequence represents a bacterial motilin receptor polypeptide.
CC The specification describes a unique sequence present in exon 1 of
CC the dog motilin receptor, which is not present in human or Sphaeroides
CC nephelus 75E7 motilin receptor sequences. The unique nucleic acid
CC sequence is useful for measuring the ability of a compound to affect
CC motilin receptor activity. Motilin receptor polynucleotides and
CC polypeptides are used to identify therapeutic compounds which are
CC useful for treating gastrointestinal diseases and disorders such as
CC gastric motility disorders, gastroparesis, irritable bowel syndrome,
CC and diarrhoea.
XX Sequence 363 AA;
SQ Query Match 33.0%; Score 673.5; DB 22; Length 363;
Best Local Similarity 49.6%; Pred. No. 3.2e-51;
Matches 135; Conservative 39; Mismatches 55; Indels 43; Gaps 4;

QY 31 SPFPGALVPVAVCLFLVGVGNGVTVMLIGRYDMRTTNNLYLGSMAVSDDLILLG 90
DB 31 SLFPASTLPVTVICILIFVVGVTGMTLLIQFKDKMTTNLYLSMAVSDDLIVPLC 90

QY 91 LPFDLYRLWRSRPWFVGPGLLCRLSLYVGGCTGYATLLHMTALSVERYLAICRPLRARVLV 150
DB 91 LPFDLYRLWKYPWLPFGVAECRLHYIPGCTSAITLHTALSISERYLAISFPLRSKVMV 150

QY 151 TRRRVALLAVLWALLSAGPFLFLVGVGEQDGISSVVVGLNGTARIASSPLASSPPULWL 210
DB 151 TRRRVQIITLAWCFALVGAAPTFLVGVGEYO-----NET----- 185

Matches	Conservative	46;	Mismatches	87;	Indels	58;	Gaps	8;
QY	5	WNGSDGEA-----REPPPALPPCD---ERRCSPFPGALVPTAVTACLCFLFVGVS	GN	56				
		: :		:		: : :		
Db	2	WNATPSEEPGFNLTLADLDWDASPGNDSLGDLLQLFPAPLAGVTATCTVALFVVGIA	GN	61				
QY	57	VVTVMILIGRYDRMRTTNLYLGMVAVSDDLILGLPFDLRLWRSPWPFGPLLCRLUS	L	116				
		: : : :		: :		: : :		
Db	62	LLTMLVVSFRPRELRTTNLYLSMAFSDDLIFLCMPDLVLWQYRPNWFNQDILLCKLF	QF	121				
QY	117	VGRGCVATLLHMTALTASYERYLAICPLRAVLVTRRRVRALITAVLMWALLSAGP	PLFL	176				
		: : :		: : :		: : :		
Db	122	VSESCTATVLTITALSVERYFAICPLRAKVVVTKGRVLIVFTIMAWAFCSAGPI	FVL	181				
QY	177	VGVEQDPGISVWPLNGTARIASSPLASSPPLWSRAPPPSGPGPETAAALFSRECR		236				
		: :		:		:		
Db	182	VGVEHE-----NGT-----DP-W-----DTNECR		199				
QY	237	PS--PAQLGALRYMLWTTAYFFLPFLCLSLYGLIGRELWSSRRPLRGPAASGRERGH	R	294				
		: : :		: : :		: : :		
Db	200	PTEFAVRSGLLTVMVWSSIFFFLPVFCLTVLYSLGRKLWRRRRGGDAVVGASLRQDNH	K	259				
QY	295	QTIVRLRWKSRRSKGAKCLOSAPPGTATQLGPLPLL		330				
		: : :		: : :		: : :		
Db	260	QTVKMIL-----GGSORALRLSIAGPILSLCLCLPSL		289				

RESULT 2

US-09-077-674-10
 / Sequence 10, Application US/09077674
 / Patent No. 6531314
 / GENERAL INFORMATION:
 / APPLICANT: Arena, Joseph P.
 / APPLICANT: Cully, Doris F.
 / APPLICANT: Feigner, Scott D.
 / APPLICANT: Howard, Andrew D.
 / APPLICANT: Liberator, Paul A.
 / APPLICANT: Schaeffer, James M.
 / APPLICANT: Van Der Ploeg, Leonardus
 / TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR FAMILY
 / NUMBER OF SEQUENCES: 16
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: Merck & Co., Inc.
 / STREET: P.O. Box 2000, 126 E. Lincoln Ave.
 / CITY: Rahway
 / STATE: NJ
 / COUNTRY: USA
 / ZIP: 07065-0900
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: Diskette
 / COMPUTER: IBM Compatible
 / OPERATING SYSTEM: DOS
 / SOFTWARE: FastSeq for Windows Version 2.0
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/09/077,674
 / FILING DATE: 3-JUN-1998
 / CLASSIFICATION: 536
 / PRIOR APPLICATION DATA:
 / APPLICATION NUMBER:
 / FILING DATE:
 / ATTORNEY/AGENT INFORMATION:
 / NAME: Cocuzzo, Anna L.
 / REGISTRATION NUMBER: 42,452
 / REFERENCE/DOCKET NUMBER: 19589P
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: 732-594-1273
 / TELEFAX: 732-594-4720
 / TELEX:
 / INFORMATION FOR SEQ ID NO: 10:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 289 amino acids
 / TYPE: amino acid
 / STRANDEDNESS: single

```

; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-077-674-10

Query Match          31.8%; Score 649; DB 4; Length 289;
Best Local Similarity 43.2%; Pred. No. 2.6e-46;
Matches 145; Conservative 46; Mismatches 87; Indels 5

QY      5 WNGSDGEGA-----REPPWALPPCD---ERRCSPFPLGALVPVTAVCL
      ||      :      ||      ||      ||      ||      ||
Db      2 WNAIPSEEPFNLTADLDWDASPGNDSLGDELILQLFPAPLLAGVTATCAV
      ||      :      ||      ||      ||      ||      ||
QY     57 VVTVMILIGRYDRMTTNLYLGSMAVSDLLILGLPFDLYRLWRSRPWVF
      ||      :      ||      ||      ||      ||      ||
Db     62 LLTLMVVSFRFRETNTNLYLSSMAFSDLLIFLCMLDLDLWQYRPWNVF
      ||      :      ||      ||      ||      ||      ||
QY    117 VGEGCTYATLLHMTALSVERYLAI CRPLRARVLVTRRRVRRLIALVLMVAVAL
      ||      :      ||      ||      ||      ||      ||
Db    122 VSESCITYATVLTITALSVERYFAICFPFLRAKVVTVKGRVLVI FVIWAVAF
      ||      :      ||      ||      ||      ||      ||
QY    177 VGEVDQPGISVVDPLNGTARIASSPLASSPPLWLSRAPPPSPGSPETAEE
      ||      :      ||      ||      ||      ||      ||
Db    182 VGEVHEB-----NGT-----DP--W-----
      ||      :      ||      ||      ||      ||      ||
QY    237 PS--PAQLGALRYMLVWTTAYFFLPFLCLSTLYGLIGELWSSRRRLRGP
      ||      :      ||      ||      ||      ||      ||
Db    200 PTEFAVRSGLLTVWVWVSSITFFFLPVPFCLTVLVSILGRKLWRRRGDAVAV
      ||      :      ||      ||      ||      ||      ||
QY    295 QTVRLRLKWSRRSGSKDACIQSAPPGTATQITGLPPLL 330
      ||      :      ||      ||      ||      ||      ||
Db    260 QTVKML-----GGSORALRLSLAGPITLSLCPLPSL 289
      ||      :      ||      ||      ||      ||      ||

```

RESULT 3

US-09-077-675A-13
Sequence 13, Application US/09077675A
Patent No. 6242199
GENERAL INFORMATION:
APPLICANT: Pai, Lee-Yuh
APPLICANT: Feighner, Scott C.
APPLICANT: Howard, Andrew D.
APPLICANT: Hong, Sheng-Shung
APPLICANT: Van Der Ploeg, Leonardus H.T.
TITLE OF INVENTION: RECEPTOR ASSAY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSQL for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/077,675A
FILING DATE: 3-JUN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19590P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEO ID NO: 13:

```
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 366 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;
US-09-077-675A-13

Query Match      31.5%; Score 643; DB 3; Length 366;
Best Local Similarity 45.1%; Pred. No. 1.1e-45;
Matches 138; Conservative 43; Mismatches 73; Indels 52; Gaps 7;

QY 5 WNGSDGPEGA-----REPPWALPPCD---ERRCSPPFLGALVPVTAVCLCLFVVGVSNG 56
DB 2 WNATPSEEPGFNLTLADLDWDASPGNDSLGLDQLLPAPLLAGVTATCVALLFVVGIAGN 61
QY 57 VVTVMILIGRYDRMTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPMWVFGPPLLCRLSLY 116
DB 62 LITMLVVSFRERLRTTNNLYLSSMAFSDDLIFLCMLDLVRLWQYRPWNFGDLLCKLQF 121
QY 117 VBGCTYATLLHMTALSVERYLAICRPLRARVLVTRRRVRLIALIWLAVALLSAGPFLFL 176
DB 122 VBSCTYATVLTITALSVERYFAICPLRAKVVVTKGRVKLVIFVIWAVAFCSAGPIFVL 181
QY 177 VGVQDPGIVSVPLNGTARIASSPLPWLRSRPPSPSGPETAEAAALFSRRCR 236
DB 182 VGVHEH-----NGT-----DP--W-----DTNECR 199
QY 237 PS--PAQLGALRVMLVTTAYFPFLPCLCLSLYGLIGRELWSSRRPLRGPASGRGRHR 294
DB 200 PTEFAVRSGLLTVMWVSSIFFLPVFCULTVLYSLIGRKLWRRRRGDVVGASLRDQNHK 259
QY 295 QTVRVL 300
DB 260 QTVKML 265

RESULT 4
US-09-077-674-13
; Sequence 13, Application US/09077674
; Patent No. 6531314
; GENERAL INFORMATION:
;   APPLICANT: Arena, Joseph P.
;   APPLICANT: Cully, Doris F.
;   APPLICANT: Feighner, Scott D.
;   APPLICANT: Howard, Andrew D.
;   APPLICANT: Liberator, Paul A.
;   APPLICANT: Schaeffer, James M.
;   APPLICANT: Van Der Ploeg, Leonardus
;   TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR FAMILY
;   NUMBER OF SEQUENCES: 16
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Merck & Co., Inc.
;     STREET: P.O. Box 2000, 126 E. Lincoln Ave.
;     CITY: Rahway
;     STATE: NJ
;     COUNTRY: USA
;     ZIP: 07065-0900
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Diskette
;     COMPUTER: IBM Compatible
;     OPERATING SYSTEM: DOS
;     SOFTWARE: FastSeq for Windows Version 2.0
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/09/077,674
;     FILING DATE: 3-JUN-1998
;     CLASSIFICATION: 536
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:
;       FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Cocuzzo, Anna L.
;     REGISTRATION NUMBER: 42,452
```

```
; REFERENCE/DOCKET NUMBER: 19589P
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 732-594-1273
;   TELEFAX: 732-594-4720
;   TELEX:
;   INFORMATION FOR SEQ ID NO: 13:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 366 amino acids
;       TYPE: amino acid
;       STRANDEDNESS: single
;       TOPOLOGY: linear
;       MOLECULE TYPE: protein
;
US-09-077-674-13

Query Match      31.5%; Score 643; DB 4; Length 366;
Best Local Similarity 45.1%; Pred. No. 1.1e-45;
Matches 138; Conservative 43; Mismatches 73; Indels 52; Gaps 7;

QY 5 WNGSDGPEGA-----REPPWALPPCD---ERRCSPPFLGALVPVTAVCLCLFVVGVSNG 56
DB 2 WNATPSEEPGFNLTLADLDWDASPGNDSLGLDQLLPAPLLAGVTATCVALLFVVGIAGN 61
QY 57 VVTVMILIGRYDRMTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPMWVFGPPLLCRLSLY 116
DB 62 LITMLVVSFRERLRTTNNLYLSSMAFSDDLIFLCMLDLVRLWQYRPWNFGDLLCKLQF 121
QY 117 VBGCTYATLLHMTALSVERYLAICRPLRARVLVTRRRVRLIALIWLAVALLSAGPFLFL 176
DB 122 VBSCTYATVLTITALSVERYFAICPLRAKVVVTKGRVKLVIFVIWAVAFCSAGPIFVL 181
QY 177 VGVQDPGIVSVPLNGTARIASSPLPWLRSRPPSPSGPETAEAAALFSRRCR 236
DB 182 VGVHEH-----NGT-----DP--W-----DTNECR 199
QY 237 PS--PAQLGALRVMLVTTAYFPFLPCLCLSLYGLIGRELWSSRRPLRGPASGRGRHR 294
DB 200 PTEFAVRSGLLTVMWVSSIFFLPVFCULTVLYSLIGRKLWRRRRGDVVGASLRDQNHK 259
QY 295 QTVRVL 300
DB 260 QTVKML 265

RESULT 5
US-09-170-496D-88
; Sequence 88, Application US/09170496D
; Patent No. 6555339
; GENERAL INFORMATION:
;   APPLICANT: Behan, Dominic P.
;   APPLICANT: Chalmers, Derek T.
;   APPLICANT: Liaw, Chen W.
;   TITLE OF INVENTION: No. 6555339-Endogenous, Constitutively Activated Human G Protein-
;   TITLE OF INVENTION: Receptors
;   FILE REFERENCE: AREN-0040
;   CURRENT APPLICATION NUMBER: US/09/170,496D
;   CURRENT FILING DATE: 1998-10-13
;   NUMBER OF SEQ ID NOS: 294
;   SOFTWARE: Patent in version 3.1
;   SEQ ID NO 88
;   LENGTH: 366
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;
US-09-170-496D-88

Query Match      31.5%; Score 643; DB 4; Length 366;
Best Local Similarity 45.1%; Pred. No. 1.1e-45;
Matches 138; Conservative 43; Mismatches 73; Indels 52; Gaps 7;

QY 5 WNGSDGPEGA-----REPPWALPPCD---ERRCSPPFLGALVPVTAVCLCLFVVGVSNG 56
DB 2 WNATPSEEPGFNLTLADLDWDASPGNDSLGLDQLLPAPLLAGVTATCVALLFVVGIAGN 61
QY 57 VVTVMILIGRYDRMTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPMWVFGPPLLCRLSLY 116
```

Db 62 LLMVLSRFRRLTNNLYLSSMAFSDLLIFLCMPDLVLWQYRWNFGDLCKLFQF 121
QY 117 VVGCTYATLLHMTALSVRYLAICRPLRARVLVTRRRVRLAVLWAVALLSAGPFLFL 176
Db 122 VSSCTYATVITLITLVSRYFAICFPLRAKVVTKGRVKLVIFVIWAVAFCSAGPIFVL 181
QY 177 VGVQDPGISVVPGLNGTARIASSPLSLRAPPSPSPGPTAEAAALFSGRECR 236
Db 182 VGVHEH-----NGT-----DP--W-----DINECR 199
QY 237 PS--PAQGLARVLMVWTTAYFFLPFLCLSLYGLIGRELWSSRRPLRGPAASGRGRH 294
Db 200 PTFEFAVRSGLLTVMVWSSIFFLPVFLVCLVLYSLIGRKLWRRRGDAVVGASLRDQNHK 259
QY 295 QTVRVL 300
Db 260 QTVKML 265

RESULT 6

US-09-077-675A-5
; Sequence 5, Application US/09077675A
; Patent No. 6242199
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 289 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-077-675A-5

Query Match 31.4%; Score 641; DB 3; Length 289;

Best Local Similarity 43.3%; Pred. No. 1.2e-45;
Matches 142; Conservative 45; Mismatches 77; Indels 64; Gaps 9;

QY 5 WNGSDGPEGA-----REPPWPAALPPCD---ERRCSPPPLGALVPVTAACLCLFVWGVSGN 56

Db 2 WNAATSEEPGNLTLFDLGDWAPPENDSLVEELLPLFPPLLAGVTATCVALFVVGIAGN 61
QY 57 VVTVMILGYRDMRTTNNLYLSSMAVSDLLILGLPFDLYRLWRSRPMVFGPILCLSLY 116
Db 62 LLMVLSRFRRLTNNLYLSSMAFSELLIFLCMPLEFLWQYRWNFGDLCKLFQF 121
QY 117 VVGCTYATLLHMTALSVRYLAICRPLRARVLVTRRRVRLAVLWAVALLSAGPFLFL 176
Db 122 VSSCTYATVITLITLVSRYFAICFPLRAKVVTKGRVKLVIFVIWAVAFCSAGPIFVL 181
QY 177 VGVQDPGISVVPGLNGTARIASSPLSLRAPPSPSPGPTAEAAALFSGRECR 236
Db 182 VGVHEH-----NGT-----DPRD-----DINECR 199
QY 237 PS--PAQGLARVLMVWTTAYFFLPFLCLSLYGLIGRELWSSRRPLRGPAASG---RER 291
Db 200 ATEFAVRSGLLTVMVWSSIFFLPVFLVCLVLYSLIGRKLW---RRKGEAAVGSRLRDQ 256
QY 292 GHRQTVRLRKWSRRGSKDACLSQAPPG 319
Db 257 NHQTVKML-----GGSQCALELSLPG 278

RESULT 7

US-09-077-674-5
; Sequence 5, Application US/09077674
; Patent No. 6531314
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR FAMILY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,674
; FILING DATE: 3-JUN-1998
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19589P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 289 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-077-674-5

```
Query Match 31.4%; Score 641; DB 4; Length 289;
Best Local Similarity 43.3%; Pred. No. 1.2e-45;
Molecule Type: protein
Matches 142; Conservative 45; Mismatches 77; Indels 64; Gaps 9;

QY 5 WNGSDGEGA-----REPPALPPCD---ERRCSPEPLGALVPVTAVALCLFVVGSGN 56
DB 2 WNATPSEEPGNUTLPDLGWDAPEPNDLSLVEELLPLFPFTPLLAGVTATCVALFVVGIGN 61
QY 57 VVTVMILIGRYDRMRTTNLYLGSMAVSDDLILGLPDLVRLWRSRPWFGLLCRLSLY 116
DB 62 LLTMLVRSRFRMRTTNLYLSSMAFSELLIFLCMPLEFLWQYRPNWNLNLLCKLQF 121
QY 117 VGGCTYATLHMTALSVERYLAICRPLARVLVTRRRVRLIAVLWAVALLSAGPLFL 176
DB 122 VSESCTYATLITLALSVERYFAICFPPLAKVVTGKRVKLVILVIAVAFCSAGPIFVL 181
QY 177 VGVQDPGIVSVVPLNGCTARIASSPLASSPPLWLSRAPPSPPSGPTAAALFSRECR 236
DB 182 VGVHEH-----NGT-----DPRD-----TNECR 199
QY 237 PS--PAQGLARVLMVTTAYFFLPFLCLSLYGLIGRELWSSRRPLRGPAASG---RER 291
DB 200 ATEFAVRSGLLTVMVWSSVFFFLPVFCLVLSLIGRKLW---RRKGEAAVGSLLRQ 256
QY 292 GHRQTVRLKRWRSRSGKDACLOSAPG 319
DB 257 NHQKQTKVML-----GGSQCALELSLPG 278

RESULT 8
US-09-719-675A-16
; Sequence 16, Application US/09077675A
; Patent No. 6243199
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ FOR WINDOWS Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 364 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-077-675A-16

Query Match 31.4%; Score 641; DB 3; Length 364;
Best Local Similarity 45.8%; Pred. No. 1.5e-45;
Matches 142; Conservative 41; Mismatches 65; Indels 62; Gaps 10;

QY 5 WNGSDGEGAPEP-----PWPALPPCD---ERRCSPPFLGALVPVTAVALCLFVVGSG 55
DB 2 WNAT--PSEEPENVTLDLDWDASPCNDLSLPDELLPLFPAPLLAGVTATCVALFVVGISG 59
QY 56 NVVTVMILIGRYDRMRTTNLYLGSMAVSDDLILGLPDLVRLWRSRPWFGLLCRLSL 115
DB 60 NLLTMLVRSRFRMRTTNLYLSSMAFSDLLIFLCMPDLVRLWQYRPNWNLNLLCKLQF 119
QY 116 YVGGCTYATLHMTALSVERYLAICRPLARVLVTRRRVRLIAVLWAVALLSAGPLFL 175
DB 120 VSESCTYATLITLALSVERYFAICFPPLAKVVTGKRVKLVILVIAVAFCSAGPIFV 179
QY 176 LVGVQDPGIVSVVPLNGTARIASSPLASSPPLWLSRAPPSPPSGPTAAALFSRECR 235
DB 180 LVGVHE-----NGT-----DPRD-----TNEC 197
QY 236 RPS--PAQGLARVLMVTTAYFFLPFLCLSLYGLIGRELWSSRRPLRGPAASG---RE 290
DB 198 RATEFAVRSGLLTVMVWSSVFFFLPVFCLVLSLIGRKLW---RR--RGDAAVGASLRD 253
QY 291 RGHRTVRLV 300
DB 254 QNHQKQTKVML 263

RESULT 9
US-09-077-674-16
; Sequence 16, Application US/09077674
; Patent No. 6531314
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR FAMILY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ FOR WINDOWS Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,674
; FILING DATE: 3-JUN-1998
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19589P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
```

```
;
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 364 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-719-485-5

Query Match      31.4%; Score 641; DB 4; Length 364;
Best Local Similarity 45.8%; Pred. No. 1.5e-45;
Matches 142; Conservative 41; Mismatches 65; Indels 62; Gaps 10;

QY 5 WNGSDGEGAREP-----DKPALPPCD---ERRCSFPLGALVPVAVCLCLFVVGVSG 55
Db 2 WNAT--PSEPEPVNTLDLWDASPGNDSLDELPLFPAPLAGVATATCVAVFVVGISG 59
QY 56 NVVTMLIGRYDMRTTNNLYLGSMAVSDLLILGLPDLRLWRSRPWVFGPLLCRLSL 115
Db 60 NLLTMLVVSFRRLTNNLYLGSMAVSDLLILGLPDLRLWRSRPWVFGPLLCRLSL 119
QY 116 YVGEQCTATLLHMTALSVERYLAICRPLARVLVTRRRVRLIAVLWAVALLSAGPFLF 175
Db 120 FVSECTYATVLTITALSVERYFAICPLRAKVVTGRVKLVILVIWAVAFCSAGPIFV 179
QY 176 LVGVEQDPGIVSVVPLNGTARIASSPLASSPPLWLSRAPPPSPGPETAEEAALFSREC 235
Db 180 LVGVEH-----NGT-----DPRD-----TNEC 197
QY 236 RPS--PAQLGALRVMLVTTTAYFPLFCLSLYGLIGRELMSRRPLRGPAASG---RE 290
Db 198 RATEFAVRSGLLTVMWVSSVFFPLFVCLTVLSLIGRKLW--RR--RGDAVAGSLRD 253
QY 291 RGRHQTVRVL 300
Db 254 QNHQKTVKML 263

RESULT 10
US-09-719-485-5
; Sequence 3, Application US/09077675A
; Patent No. 6242199
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.

;
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 353 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-077-675A-3

Query Match      31.3%; Score 639.5; DB 3; Length 353;
Best Local Similarity 47.2%; Pred. No. 2e-45;
Matches 137; Conservative 39; Mismatches 61; Indels 53; Gaps 7;

QY 19 WPAIPPCD---ERRCSFPLGALVPVAVCLCLFVVGVSGNVVTVMLIGRYDMRTTNNL 75
Db 8 WDAPPENDSLVEELLPLFPPLLAGVATATCVAVFVVGIAAGNLLTMLVVSFRFRRTTNNL 67
QY 76 YLGSMAVSDLLILGLPDLRLWRSRPWVFGPLLCRLSLYVGEQCTATLLHMTALSVE 135
Db 68 YLGSMAVSDLLILGLPDLRLWRSRPWVFGPLLCRLSLYVGEQCTATLLHMTALSVE 127
QY 136 RYLAICRPLARVLVTRRRVRLIAVLWAVALLSAGPFLFVGVQDPGIVSVVPLNGTA 195
Db 128 RYFAICPLRAKVVTGRVKLVILVIWAVAFCSAGPIFVLVGVGHEH-----NGT- 177
QY 196 RIASSPLASSPPLWLSRAPPPSPGPETAEEAALFSRECRPS--PAQLGALRVMLVTT 253
Db 178 -----DPRD-----TNECRATEFAVRSGLLTVMWVSS 205
QY 254 AVFPLFCLSLYGLIGRELMSRRPLRGPAASG---RERGRHQTVRVL 300
Db 206 VFFFLPVCLTVLSLIGRKLW---RRKGEAAVSGSLRDQNHQKTVKML 252

RESULT 11
US-09-077-674-3
; Sequence 3, Application US/09077674
; Patent No. 6531314
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR FAMILY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,674
; FILING DATE: 3-JUN-1998
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
```



```
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19589P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 353 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-077-674-3

Query Match 31.3%; Score 639.5; DB 4; Length 353;
Best Local Similarity 47.2%; Pred. No. 28-45;
Matches 137; Conservative 39; Mismatches 61; Indels 53; Gaps 7;

QY 19 WPAIPDCD---ERRCSPPFLGALVPVTAVCLCLFVVGSGNVVTVMILIGRYDRMTITNL 75
Db 8 WDAPPENDSLVEELLPLFPPTPLLAGVTATCVAFVVGAGNLLTMLVVSFRERMTITNL 67
QY 76 YLGSMVSDLLILGLPFDLYLRWSRPVWFGPCLLRLSLYVGEGETYATLLHMTALSVE 135
Db 68 YLSSMAFSDLLIFLCMPDLFLRLQVPRWNLGNLLCKLFQVSESCYATVLTITALSVE 127
QY 136 RYLAICRPLARVLTTRRRVRLAIVLWAVALLSAGPFLVGVQDPGIVSVVGLNGTA 195
Db 128 RYFAICFPFLRAKVVTGKRVKLVILVIMAVAFCSAGPIFVLVGEHD-----NGT- 177
QY 196 RIASSPLASSPPLWLSRAPPPSPGPETAEEAALFSRCRPS--PAQLGALRVMLWVTT 253
Db 178 -----DPRD-----TNECRATEFAVRSGLLIVMWVSS 205
QY 254 AYFFLPFLCLSLYGLIGRELMSRRPLRGPASG---RRGRHQTVRVL 300
Db 206 VPFPLPFLCLTVLYSLIGRLKLM---RRKRGAAVGSRLRDQNHKQTVKML 252

RESULT 12
US-09-077-675A-8
; Sequence 8, Application US/09077675A
; Patent No. 6242199
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:

; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 361 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-077-675A-8

Query Match 31.3%; Score 639.5; DB 3; Length 361;
Best Local Similarity 47.0%; Pred. No. 28-45;
Matches 135; Conservative 41; Mismatches 64; Indels 47; Gaps 6;

QY 19 WPAIPDCD---ERRCSPPFLGALVPVTAVCLCLFVVGSGNVVTVMILIGRYDRMTITNL 75
Db 16 WDASPNDLSLDELLOLFPAPLLAGVTATCVAFVVGAGNLLTMLVVSFRERMTITNL 75
QY 76 YLGSMVSDLLILGLPFDLYLRWSRPVWFGPCLLRLSLYVGEGETYATLLHMTALSVE 135
Db 76 YLSSMAFSDLLIFLCMPDLFLRLQVPRWNLGNLLCKLFQVSESCYATVLTITALSVE 135
QY 136 RYLAICRPLARVLTTRRRVRLAIVLWAVALLSAGPFLVGVQDPGIVSVVGLNGTA 195
Db 136 RYFAICFPFLRAKVVTGKRVKLVILVIMAVAFCSAGPIFVLVGEHE-----NGT- 185
QY 196 RIASSPLASSPPLWLSRAPPPSPGPETAEEAALFSRCRPS--PAQLGALRVMLWVTT 253
Db 186 -----DP--W-----DTNECRTEFAVRSGLLIVMWVSS 213
QY 254 AYFFLPFLCLSLYGLIGRELMSRRPLRGPASGRGRHQTVRVL 300
Db 214 IFFFLPFLCLTVLYSLIGRLKLMRRRRRGDAVVGASLRDQNHKQTVKML 260

RESULT 13
US-09-077-674-8
; Sequence 8, Application US/09077674
; Patent No. 6531314
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR FAMILY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,674
; FILING DATE: 3-JUN-1998
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
```

```

; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19589P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 361 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-077-674-8

Query Match 31.3%; Score 639.5; DB 4; Length 361;
Best Local Similarity 47.0%; Pred. No. 2e-45;
Matches 135; Conservative 41; Mismatches 64; Indels 47; Gaps 6;

QY 19 WPAIPPCD---ERRCSPPGALVPVTAVALCLFVGVSGNVVTVMLIGRYDRMTTNL 75
DB 16 WDASPGNDSLGDELQLFPAPLAGVATATCVAFVVGVIAGNLLTMLVSRFRELRTTNL 75
QY 76 YLGSMVSDLLILGLPDLVRLWRSRPPVGPPLCLSLYVGGCTYATLLHMTALSVE 135
DB 76 YLSMAFSDLLIFLCMPDLVRLWQYRPWFGDILLCKLFQFVSESCYATVLTITALSVE 135
QY 136 RYLAICPLRARLVTRRRVRLIALVWALLSAGPFLVGVQDPGISVVPGLNGTA 195
DB 136 RYPAICFPLRAKVTVTKRVKLVIFVWAVAFCSAGPIFVLVGEHE-----NGT- 185
QY 196 RIASSPLASSPPLWLSRAPPPSPGPETAFAALFSRECRPS--PAQLGALRVMLVTT 253
DB 186 -----DP-W-----DTNECRPTEFAVRSGLLTVMVWVSS 213
QY 254 AYFPLPCLSLYGLIGRLWSSRPLRGPAASGRGRHQTQTVRL 300
DB 214 IFPLPVLTVLSLIGRLWRRRRRGDAVVGASLRDQNHKQTVKML 260

RESULT 14
US-09-170-496D-210
; Sequence 210, Application US/09170496D
; Patent No. 6553339
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. 6553339-Endogenous, Constitutively Activated Human G Protein-
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/09/170,496D
; CURRENT FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 210
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-170-496D-210

Query Match 31.2%; Score 637; DB 4; Length 366;
Best Local Similarity 44.8%; Pred. No. 3.3e-45;
Matches 137; Conservative 43; Mismatches 74; Indels 52; Gaps 7;

QY 5 WNGSDGPEGA-----REPPWPAIPPCD---ERRCSPPGALVPVTAVALCLFVGVSGN 56
DB 2 WNAIPSEPGNLTALDLDWDGSGNDSLGDELQLFPAPLAGVATATCVAFVVGVIAGN 61

```

```

QY 57 VVTMLIGRYDRMTTNLNLVYLGSMVSDLLILGLPDLVRLWRSRPPVGPPLCLSLY 116
DB 62 LLTMLVSRFRELRTTNLNLVYLSMAFSDLLIFLCMPDLVRLWQYRPWFGDILLCKLFQF 121
QY 117 VGGCTYATLLHMTALSVERYLAI CRPLRARLVTRRRVRLIALVWALLSAGPFL 176
DB 122 VSSCTYATVLTITALSVERYFAICFPLRAKVTVTKRVKLVIFVWAVAFCSAGPIFVL 181
QY 177 VGVQDPGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPETAFAALFSRECR 236
DB 182 VGVGEHE-----NGT-----DP-W-----DTNECR 199
QY 237 PS--PAQLGALRVMLVTTAYFPLPCLSLYGLIGRLWSSRPLRGPAASGRGRH 294
DB 200 PTEFAVRSGLLTVMVWVSSIFPLPVLVCLTVLSLIGRLWRRRRRGDAVVGASLRDQNHK 259
QY 295 QTVRVL 300
DB 260 QTKKML 265

RESULT 15
US-09-077-675A-7
; Sequence 7, Application US/09077675A
; Patent No. 6242199
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 302 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-077-675A-7

Query Match 27.1%; Score 554; DB 3; Length 302;
Best Local Similarity 47.7%; Pred. No. 2e-38;
Matches 116; Conservative 35; Mismatches 48; Indels 44; Gaps 5;

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 1, 2004, 06:23:05 ; Search time 141.727 Seconds
(without alignments)
546.456 Million cell updates/sec

Title: US-09-719-485-5

Perfect score: 2043

Sequence: 1 MGSPWNGSDGPGAREPPMP.....WQNLHKGHFADVDLLSVL 386

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 733937 seqs, 200641211 residues

Total number of hits satisfying chosen parameters: 733937

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA:

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10F_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US10G_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1581	77.4	412	12	US-10-318-661-28
2	1581	77.4	412	12	US-10-206-677-2
3	1581	77.4	412	12	US-10-225-567A-473
4	1581	77.4	412	15	US-10-290-078-15
5	1575	77.1	412	12	US-10-417-820A-130
6	1575	77.1	412	12	US-10-417-820A-150
7	649	31.8	289	12	US-10-303-204A-10
8	649	31.8	289	15	US-10-225-567A-140
9	643	31.5	366	12	US-10-303-204A-13
10	643	31.5	366	12	US-10-251-385-88
11	641	31.4	289	12	US-10-303-204A-5
12	641	31.4	384	12	US-10-303-204A-16
13	639.5	31.3	353	12	US-10-303-204A-3
14	639.5	31.3	361	12	US-10-303-204A-8
15	637	31.2	366	15	US-10-251-385-210

16	554	27.1	302	12	US-10-303-204A-7	Sequence 7, Appli
17	552	27.0	302	12	US-10-303-204A-2	Sequence 2, Appli
18	454	22.2	271	12	US-10-303-204A-12	Sequence 12, Appli
19	355.5	17.4	445	12	US-10-240-145-53	Sequence 53, Appli
20	355.5	17.4	445	12	US-10-240-145-139	Sequence 139, Appli
21	354.5	17.4	403	12	US-10-353-690-10	Sequence 10, Appli
22	354.5	17.4	403	15	US-10-251-385-114	Sequence 114, Appli
23	354.5	17.4	403	15	US-10-225-567A-540	Sequence 540, Appli
24	354.5	17.4	403	15	US-10-290-078-18	Sequence 18, Appli
25	352.5	17.3	403	15	US-10-251-385-224	Sequence 224, Appli
26	335	16.4	412	15	US-10-225-567A-557	Sequence 557, Appli
27	335	16.4	415	12	US-10-272-983-12	Sequence 12, Appli
28	335	16.4	415	12	US-10-393-807-12	Sequence 12, Appli
29	335	16.4	415	12	US-10-417-820A-12	Sequence 12, Appli
30	330	16.2	418	15	US-10-225-567A-207	Sequence 207, Appli
31	326.5	16.0	418	12	US-09-826-509-535	Sequence 535, Appli
32	322	15.8	249	11	US-09-782-974C-18	Sequence 18, Appli
33	309	15.1	419	9	US-09-804-551B-26	Sequence 26, Appli
34	309	15.1	428	15	US-10-270-333-114	Sequence 114, Appli
35	285.5	14.0	418	12	US-10-369-493-5319	Sequence 5319, Ap
36	274.5	13.4	410	12	US-09-826-509-537	Sequence 537, Appli
37	274.5	13.4	410	15	US-10-225-567A-432	Sequence 432, Appli
38	268.5	13.1	416	12	US-10-205-219-21	Sequence 21, Appli
39	262	12.8	362	11	US-09-992-331-13	Sequence 13, Appli
40	262	12.8	362	11	US-09-964-923A-24	Sequence 24, Appli
41	261	12.8	595	12	US-10-314-076-17	Sequence 17, Appli
42	261	12.8	595	15	US-10-270-333-195	Sequence 195, Appli
43	259.5	12.7	367	9	US-09-823-114-23	Sequence 23, Appli
44	259.5	12.7	367	15	US-10-290-748-23	Sequence 23, Appli
45	255.5	12.5	378	12	US-10-369-493-6848	Sequence 6848, Ap

ALIGNMENTS

RESULT 1

US-10-318-661-28
; Sequence 28, Application US/10318661
; Publication No. US20030167476A1
; GENERAL INFORMATION:
; APPLICANT: Conklin, Bruce R.
; TITLE OF INVENTION: Selective Target Cell Activation By
; Publication No. US20030167476A1
; TITLE OF INVENTION: Expression of A G Protein-Coupled Receptor Activated
; TITLE OF INVENTION: Superiorly By Synthetic Ligand
; FILE REFERENCE: UCAL-049CIP2
; CURRENT APPLICATION NUMBER: US/10/318,661
; CURRENT FILING DATE: 2003-05-05
; PRIOR APPLICATION NUMBER: US 09/341,446
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US97/05334
; PRIOR FILING DATE: 1997-03-25
; PRIOR APPLICATION NUMBER: US 08/622,348
; PRIOR FILING DATE: 1996-03-26
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-318-661-28

Query Match 77.4%; Score 1581; DB 12; Length 412;
Best Local Similarity 100.0%; Pred. No. 2,4e-118;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MGSPWNGSDGPGAREPPMPALPPCDERCSPPFLGALVPVTVAVCLFVWGSNGVTV 60
Db	1	MGSPWNGSDGPGAREPPMPALPPCDERCSPPFLGALVPVTVAVCLFVWGSNGVTV 60
Qy	61	MLIGRYDRMTTNNLYLGSMAVSDLLILGLPDLRLWRSRPPVGLLRLSLVVGEG 120
Db	61	MLIGRYDRMTTNNLYLGSMAVSDLLILGLPDLRLWRSRPPVGLLRLSLVVGEG 120

QY 121 CTYATLLHMTALSVRYLAICRPLRARVLVTRRRVRALIAVLWAVALLSAGPFLVLGVGE 180
DB 121 CTYATLLHMTALSVRYLAICRPLRARVLVTRRRVRALIAVLWAVALLSAGPFLVLGVGE 180
QY 181 QDPSISVPLGNTARIASSPLASSPPLWLSRAPPPSPGPGTAAALFSRECRPSA 240
DB 181 QDPSISVPLGNTARIASSPLASSPPLWLSRAPPPSPGPGTAAALFSRECRPSA 240
QY 241 QLGALRVMLWVTYATFFLPCLCLISILYGLIGRELWSSRRPLRGPAASGRGRHQTIVRL 300
DB 241 QLGALRVMLWVTYATFFLPCLCLISILYGLIGRELWSSRRPLRGPAASGRGRHQTIVRL 300

RESULT 2

US-10-206-677-2
; Sequence 2, Application US/10206677
; Publication No. US20030186336A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences, Inc.
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burner, Glenna C.
; APPLICANT: Roush, Christine L.
; APPLICANT: Kullander, Bruce G.
; TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC COMPOSITIONS AND METHODS RELATED TO
; FILE OF INVENTION: GPCR 38, A G PROTEIN-COUPLED RECEPTOR (GPCR)
; FILE REFERENCE: 1920-1-8
; CURRENT APPLICATION NUMBER: US/10/206,677
; CURRENT FILING DATE: 2002-11-06
; PRIOR APPLICATION NUMBER: 60/250,251
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 60/250,452
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-206-677-2

Query Match 77.4%; Score 1581; DB 12; Length 412;
Best Local Similarity 100.0%; Pred. No. 2.4e-118;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGSPWNGSDGPEGAREPPMPALPPCDERRCSPPPLGALVPVTAVCLCLFVVGSGNVVTV 60
DB 1 MGSPWNGSDGPEGAREPPMPALPPCDERRCSPPPLGALVPVTAVCLCLFVVGSGNVVTV 60
QY 61 MLIGRYDRMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWFVGPPLLCRLSLYVGE 120
DB 61 MLIGRYDRMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWFVGPPLLCRLSLYVGE 120
QY 121 CTYATLLHMTALSVRYLAICRPLRARVLVTRRRVRALIAVLWAVALLSAGPFLVLGVGE 180
DB 121 CTYATLLHMTALSVRYLAICRPLRARVLVTRRRVRALIAVLWAVALLSAGPFLVLGVGE 180
QY 181 QDPSISVPLGNTARIASSPLASSPPLWLSRAPPPSPGPGTAAALFSRECRPSA 240
DB 181 QDPSISVPLGNTARIASSPLASSPPLWLSRAPPPSPGPGTAAALFSRECRPSA 240
QY 241 QLGALRVMLWVTYATFFLPCLCLISILYGLIGRELWSSRRPLRGPAASGRGRHQTIVRL 300
DB 241 QLGALRVMLWVTYATFFLPCLCLISILYGLIGRELWSSRRPLRGPAASGRGRHQTIVRL 300

RESULT 3

US-10-225-567A-473
; Sequence 473, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burner, Glenna C.

; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; PRIOR APPLICATION NUMBER: 2001-12-19
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 473
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-225-567A-473

Query Match 77.4%; Score 1581; DB 15; Length 412;
Best Local Similarity 100.0%; Pred. No. 2.4e-118;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGSPWNGSDGPEGAREPPMPALPPCDERRCSPPPLGALVPVTAVCLCLFVVGSGNVVTV 60
DB 1 MGSPWNGSDGPEGAREPPMPALPPCDERRCSPPPLGALVPVTAVCLCLFVVGSGNVVTV 60
QY 61 MLIGRYDRMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWFVGPPLLCRLSLYVGE 120
DB 61 MLIGRYDRMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWFVGPPLLCRLSLYVGE 120
QY 121 CTYATLLHMTALSVRYLAICRPLRARVLVTRRRVRALIAVLWAVALLSAGPFLVLGVGE 180
DB 121 CTYATLLHMTALSVRYLAICRPLRARVLVTRRRVRALIAVLWAVALLSAGPFLVLGVGE 180
QY 181 QDPSISVPLGNTARIASSPLASSPPLWLSRAPPPSPGPGTAAALFSRECRPSA 240
DB 181 QDPSISVPLGNTARIASSPLASSPPLWLSRAPPPSPGPGTAAALFSRECRPSA 240
QY 241 QLGALRVMLWVTYATFFLPCLCLISILYGLIGRELWSSRRPLRGPAASGRGRHQTIVRL 300
DB 241 QLGALRVMLWVTYATFFLPCLCLISILYGLIGRELWSSRRPLRGPAASGRGRHQTIVRL 300

RESULT 4

US-10-290-078-15
; Sequence 15, Application US/10290078
; Publication No. US20030124596A1
; GENERAL INFORMATION:
; APPLICANT: Carroll, Joseph A.
; TITLE OF INVENTION: Methods and Compositions for Treating
; TITLE OF INVENTION: Hematological Disorders Using 232, 2059, 10630, 12848, 13875,
; FILE OF INVENTION: 14395, 14618, 17692 or 58874
; FILE REFERENCE: MPI2001-288PI(M)
; CURRENT APPLICATION NUMBER: US/10/290,078
; CURRENT FILING DATE: 2002-11-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-290-078-15

Query Match 77.4%; Score 1581; DB 15; Length 412;
Best Local Similarity 100.0%; Pred. No. 2.4e-118;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGSPWNGSDGPEGAREPPMPALPPCDERRCSPPPLGALVPVTAVCLCLFVVGSGNVVTV 60
DB 1 MGSPWNGSDGPEGAREPPMPALPPCDERRCSPPPLGALVPVTAVCLCLFVVGSGNVVTV 60
QY 61 MLIGRYDRMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWFVGPPLLCRLSLYVGE 120
DB 61 MLIGRYDRMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWFVGPPLLCRLSLYVGE 120
QY 121 CTYATLLHMTALSVRYLAICRPLRARVLVTRRRVRALIAVLWAVALLSAGPFLVLGVGE 180

```
Db 121 CTYATLLHMTALSVERYLAICRPLRARVLTTRRRVRLIAVLWAVALLSAGPFLVLGVE 180
QY 181 QDQGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAEAAALFSRECRPSA 240
Db 181 QDQGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAEAAALFSRECRPSA 240
QY 241 QLGALRVMLWVTYAYFFLPFLCLISILYGLIGRELWSSRRPLRGPAASGRGRHQRQTVRVL 300
Db 241 QLGALRVMLWVTYAYFFLPFLCLISILYGLIGRELWSSRRPLRGPAASGRGRHQRQTVRVL 300

RESULT 5
US-10-417-820A-130
; Sequence 130, Application US/10417820A
; Publication No. US20030229216A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Ruoping
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lowitz, Kevin
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Behan, Dominic P.
; TITLE OF INVENTION: Constitutively Activated Human G Protein Coupled
; FILE REFERENCE: 7.US28.CON
; CURRENT APPLICATION NUMBER: US/10/417,820A
; PRIOR FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: 09/416,760
; PRIOR FILING DATE: 1999-10-12
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: 60/110,060
; PRIOR FILING DATE: 1998-11-27
; PRIOR APPLICATION NUMBER: 60/120,416
; PRIOR FILING DATE: 1999-02-16
; PRIOR APPLICATION NUMBER: 60/121,852
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/109,213
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/123,945
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/123,948
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/123,951
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 130
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-417-820A-130

Query Match 77.1%; Score 1575; DB 12; Length 412;
Best Local Similarity 99.7%; Pred. No. 7.4e-118;
Matches 299; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGSPWNGSDGPEGAREPPWALPPCDERRCSPPLGALVPVTAVCCLCFVVGSGNVTV 60
Db 1 MGSPWNGSDGPEGAREPPWALPPCDERRCSPPLGALVPVTAVCCLCFVVGSGNVTV 60

QY 61 MLIGRYDRMTTNNLYGSMVSDLLILLGLPFDLYRLWRSRPPWFGPLLCRLSLYVGE 120
Db 61 MLIGRYDRMTTNNLYGSMVSDLLILLGLPFDLYRLWRSRPPWFGPLLCRLSLYVGE 120

QY 121 CTYATLLHMTALSVERYLAICRPLRARVLTTRRRVRLIAVLWAVALLSAGPFLVLGVE 180
Db 121 CTYATLLHMTALSVERYLAICRPLRARVLTTRRRVRLIAVLWAVALLSAGPFLVLGVE 180

QY 181 QDQGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAEAAALFSRECRPSA 240
Db 181 QDQGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAEAAALFSRECRPSA 240

QY 241 QLGALRVMLWVTYAYFFLPFLCLISILYGLIGRELWSSRRPLRGPAASGRGRHQRQTVRVL 300
```

```
Db 181 QDQGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAEAAALFSRECRPSA 240
QY 241 QLGALRVMLWVTYAYFFLPFLCLISILYGLIGRELWSSRRPLRGPAASGRGRHQRQTVRVL 300
Db 241 QLGALRVMLWVTYAYFFLPFLCLISILYGLIGRELWSSRRPLRGPAASGRGRHQRQTVRVL 300

RESULT 6
US-10-417-820A-150
; Sequence 150, Application US/10417820A
; Publication No. US20030229216A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Ruoping
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lowitz, Kevin
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Behan, Dominic P.
; TITLE OF INVENTION: Constitutively Activated Human G Protein Coupled
; FILE REFERENCE: 7.US28.CON
; CURRENT APPLICATION NUMBER: US/10/417,820A
; PRIOR FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: 09/416,760
; PRIOR FILING DATE: 1999-10-12
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: 60/110,060
; PRIOR FILING DATE: 1998-11-27
; PRIOR APPLICATION NUMBER: 60/120,416
; PRIOR FILING DATE: 1999-02-16
; PRIOR APPLICATION NUMBER: 60/121,852
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 60/109,213
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/123,944
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/123,945
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/123,948
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/123,951
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 150
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-417-820A-150

Query Match 77.1%; Score 1575; DB 12; Length 412;
Best Local Similarity 99.7%; Pred. No. 7.4e-118;
Matches 299; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGSPWNGSDGPEGAREPPWALPPCDERRCSPPLGALVPVTAVCCLCFVVGSGNVTV 60
Db 1 MGSPWNGSDGPEGAREPPWALPPCDERRCSPPLGALVPVTAVCCLCFVVGSGNVTV 60

QY 61 MLIGRYDRMTTNNLYGSMVSDLLILLGLPFDLYRLWRSRPPWFGPLLCRLSLYVGE 120
Db 61 MLIGRYDRMTTNNLYGSMVSDLLILLGLPFDLYRLWRSRPPWFGPLLCRLSLYVGE 120

QY 121 CTYATLLHMTALSVERYLAICRPLRARVLTTRRRVRLIAVLWAVALLSAGPFLVLGVE 180
Db 121 CTYATLLHMTALSVERYLAICRPLRARVLTTRRRVRLIAVLWAVALLSAGPFLVLGVE 180

QY 181 QDQGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAEAAALFSRECRPSA 240
Db 181 QDQGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAEAAALFSRECRPSA 240

QY 241 QLGALRVMLWVTYAYFFLPFLCLISILYGLIGRELWSSRRPLRGPAASGRGRHQRQTVRVL 300
```

```

Db      241 QLGALRVMLVTTTAYFFLPCLSLYGLIGRELWSSRRPLRGPAASGRGRGHRQTKRVL 300
|||||
RESULT 7
US-10-303-204A-10
; Sequence 10, Application US/10303204A
; Publication No. US20030166144A1
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus H. T.
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR
; TITLE OF INVENTION: FAMILY
; FILE REFERENCE: 19589PCA
; CURRENT APPLICATION NUMBER: US/10/303,204A
; PRIOR FILING DATE: 2002-11-25
; PRIOR FILING DATE: 1998-06-03
; PRIOR FILING DATE: 1996-12-10
; PRIOR FILING DATE: 1996-06-06
; PRIOR FILING DATE: 1995-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 289
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-303-204A-10

Query Match      31.8%; Score 649; DB 12; Length 289;
Best Local Similarity 43.2%; Pred. No. 5.2e-44;
Matches 145; Conservative 46; Mismatches 87; Indels 58; Gaps 8;

QY      5 WNGSDGPEGA-----REPPWALPPCD---ERRCSPPFLGALVPVTAVCLCLFVVGVSNG 56
Db      2 WNAATPSEEPGFNLTLADLDWDASPGNDSLGDELLQLFPAPLLAGVTATCVALFVVGVIAGN 61
|||||
QY      57 VVTVMILIGRYDRMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWVFGPILCLSLY 116
Db      62 LLTMLVVSRELRITNNLYLSMAFSDLLIFLCMPDLVRLWQYRPWVNFGLLCKLQF 121
|||||
QY      117 VGEGCTYATLLHMTALSVERYLAICRPLARVLVTRRRVRLAIVLWAVALLSAGPFLFL 176
Db      122 VSECTYATVLTITALSVERYFAICPLRAKVVTGKRVKLVFVWVAVAFCSAGIFVL 181
|||||
QY      177 VGVQDPGIVSVPLNGTARIASSPPLSLRAPPSPSPGPETAFAALFSRECR 236
Db      182 VGVHEH-----NGT-----DP--W-----DNECR 199
|||||
QY      237 PS--PAQLGALRVMLVTTTAYFFLPCLSLYGLIGRELWSSRRPLRGPAASGRGRGHR 294
Db      200 PTEFAVRSGLLTVWVWSSIFFLPVFCITVLYSLIGRLWRRRGDAVVGASLRDQNHK 259
|||||
QY      295 QTVRLRKWSRGSKDACIQSAPPGTAQTIGPLPL 330
Db      260 QTVKML-----GGSORALRLSLAGPILSLCLPLSL 289
|||||

RESULT 9
US-10-303-204A-13
; Sequence 13, Application US/10303204A
; Publication No. US20030166144A1
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus H. T.
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR
; TITLE OF INVENTION: FAMILY
; FILE REFERENCE: 19589PCA
; CURRENT APPLICATION NUMBER: US/10/303,204A
; PRIOR FILING DATE: 2002-11-25
; PRIOR FILING DATE: 1998-06-03
; PRIOR FILING DATE: 1996-12-10
; PRIOR FILING DATE: 1996-06-06
; PRIOR FILING DATE: 1995-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13

```

```

; APPLICANT: Burmer, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR FILING DATE: 2000-12-19
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 140
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-225-567A-140

Query Match      31.8%; Score 649; DB 15; Length 289;
Best Local Similarity 43.2%; Pred. No. 5.2e-44;
Matches 145; Conservative 46; Mismatches 87; Indels 58; Gaps 8;

QY      5 WNGSDGPEGA-----REPPWALPPCD---ERRCSPPFLGALVPVTAVCLCLFVVGVSNG 56
Db      2 WNAATPSEEPGFNLTLADLDWDASPGNDSLGDELLQLFPAPLLAGVTATCVALFVVGVIAGN 61
|||||
QY      57 VVTVMILIGRYDRMRTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWVFGPILCLSLY 116
Db      62 LLTMLVVSRELRITNNLYLSMAFSDLLIFLCMPDLVRLWQYRPWVNFGLLCKLQF 121
|||||
QY      117 VGEGCTYATLLHMTALSVERYLAICRPLARVLVTRRRVRLAIVLWAVALLSAGPFLFL 176
Db      122 VSECTYATVLTITALSVERYFAICPLRAKVVTGKRVKLVFVWVAVAFCSAGIFVL 181
|||||
QY      177 VGVQDPGIVSVPLNGTARIASSPPLSLRAPPSPSPGPETAFAALFSRECR 236
Db      182 VGVHEH-----NGT-----DP--W-----DNECR 199
|||||
QY      237 PS--PAQLGALRVMLVTTTAYFFLPCLSLYGLIGRELWSSRRPLRGPAASGRGRGHR 294
Db      200 PTEFAVRSGLLTVWVWSSIFFLPVFCITVLYSLIGRLWRRRGDAVVGASLRDQNHK 259
|||||
QY      295 QTVRLRKWSRGSKDACIQSAPPGTAQTIGPLPL 330
Db      260 QTVKML-----GGSORALRLSLAGPILSLCLPLSL 289
|||||

RESULT 9
US-10-303-204A-13
; Sequence 13, Application US/10303204A
; Publication No. US20030166144A1
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus H. T.
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR
; TITLE OF INVENTION: FAMILY
; FILE REFERENCE: 19589PCA
; CURRENT APPLICATION NUMBER: US/10/303,204A
; PRIOR FILING DATE: 2002-11-25
; PRIOR FILING DATE: 1998-06-03
; PRIOR FILING DATE: 1996-12-10
; PRIOR FILING DATE: 1996-06-06
; PRIOR FILING DATE: 1995-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13

```



```

; LENGTH: 366
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-303-204A-13

Query Match      31.5%; Score 643; DB 12; Length 366;
Best Local Similarity 45.1%; Pred. No. 2.1e-43;
Matches 138; Conservative 43; Mismatches 73; Indels 52; Gaps 7;

Qy 5 WNGSDGPEGA-----REPPWALPPCD---ERRCSPPPLGALVPVTAACLCFVVGSGN 56
Db 2 WNA TPSEEPGNLTADLDWDASGNSLGDDELQLPAPLACGTATCTVALFVVGAGN 61

Qy 57 VVTVMILGRYDRMTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWFGLLCRLSLY 116
Db 62 LLTMLVSRPRELRTTNNLYLSSMAFSDLLIFLCMPDLVRLWQYRPNWFGDLCKLQF 121

Qy 117 VGECTATLLHMTALSVERYLAICRPLARVLVTRRRVRLALIAVLWALLSAGPFL 176
Db 122 VSECTATVLTITALSVERYFAICFPLRAKVVTGKRVKLVIFVIWAVAFCSAGPIFVL 181

Qy 177 VGVQDPGISVWFLNGTARIASSPLASPPMLSRAPPPSPGPGTAAALFSRECR 236
Db 182 VGVHE-----NGT-----DP--W-----DTNECR 199

Qy 237 PS--PAQGLARVLMVWTTAYFPFLCLSLYGLIGRELWSSRRPLRGPAASGRGRHR 294
Db 200 PTEFAVRSGLLTVMWVSSIFFFPVFCULTVLSLIGRKLWRRRRGDVAVGASLRDQNHK 259

Qy 295 QTVRVL 300
Db 260 QTVKML 265

RESULT 10
US-10-251-385-88
; Sequence 88, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
; TITLE OF INVENTION: Protein-Coupled
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/10/251,385
; CURRENT FILING DATE: 2002-09-20
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 88
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-251-385-88

Query Match      31.5%; Score 643; DB 15; Length 366;
Best Local Similarity 45.1%; Pred. No. 2.1e-43;
Matches 138; Conservative 43; Mismatches 73; Indels 52; Gaps 7;

Qy 5 WNGSDGPEGA-----REPPWALPPCD---ERRCSPPPLGALVPVTAACLCFVVGSGN 56
Db 2 WNA TPSEEPGNLTADLDWDASGNSLGDDELQLPAPLACGTATCTVALFVVGAGN 61

Qy 57 VVTVMILGRYDRMTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWFGLLCRLSLY 116
Db 62 LLTMLVSRPRELRTTNNLYLSSMAFSDLLIFLCMPDLVRLWQYRPNWFGDLCKLQF 121

Qy 117 VGECTATLLHMTALSVERYLAICRPLARVLVTRRRVRLALIAVLWALLSAGPFL 176
Db 122 VSECTATVLTITALSVERYFAICFPLRAKVVTGKRVKLVIFVIWAVAFCSAGPIFVL 181

```

```

Qy 177 VGVQDPGISVWFLNGTARIASSPLASPPMLSRAPPPSPGPGTAAALFSRECR 236
Db 182 VGVHE-----NGT-----DP--W-----DTNECR 199

Qy 237 PS--PAQGLARVLMVWTTAYFPFLCLSLYGLIGRELWSSRRPLRGPAASGRGRHR 294
Db 200 PTEFAVRSGLLTVMWVSSIFFFPVFCULTVLSLIGRKLWRRRRGDVAVGASLRDQNHK 259

Qy 295 QTVRVL 300
Db 260 QTVKML 265

RESULT 11
US-10-303-204A-5
; Sequence 5, Application US/10303204A
; Publication No. US20030106614A1
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus H. T.
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR
; TITLE OF INVENTION: FAMILY
; FILE REFERENCE: 19589PCA
; CURRENT APPLICATION NUMBER: US/10/303,204A
; CURRENT FILING DATE: 2002-11-25
; PRIOR APPLICATION NUMBER: 09/077,674
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: PCT/US96/19445
; PRIOR FILING DATE: 1996-12-10
; PRIOR APPLICATION NUMBER: 60/018,962
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: 60/008,582
; PRIOR FILING DATE: 1995-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 289
; TYPE: PRT
; ORGANISM: sus scrofa
US-10-303-204A-5

Query Match      31.4%; Score 641; DB 12; Length 289;
Best Local Similarity 43.3%; Pred. No. 2.3e-43;
Matches 142; Conservative 45; Mismatches 77; Indels 64; Gaps 9;

Qy 5 WNGSDGPEGA-----REPPWALPPCD---ERRCSPPPLGALVPVTAACLCFVVGSGN 56
Db 2 WNA TPSEEPGNLTADLDWDASGNSLVEELPLPPTPLLAGVTATCTVALFVVGAGN 61

Qy 57 VVTVMILGRYDRMTTNNLYLGSMAVSDLLILGLPFDLYRLWRSRPWFGLLCRLSLY 116
Db 62 LLTMLVSRPRELRTTNNLYLSSMAFSELLIFLCMPLELFLWQYRPNWNLGKLLCKLQF 121

Qy 117 VGECTATLLHMTALSVERYLAICRPLARVLVTRRRVRLALIAVLWALLSAGPFL 176
Db 122 VSECTATVLTITALSVERYFAICFPLRAKVVTGKRVKLVIFVIWAVAFCSAGPIFVL 181

Qy 177 VGVQDPGISVWFLNGTARIASSPLASPPMLSRAPPPSPGPGTAAALFSRECR 236
Db 182 VGVHE-----NGT-----DPRD-----TNECR 199

Qy 237 PS--PAQGLARVLMVWTTAYFPFLCLSLYGLIGRELWSSRRPLRGPAASG---RER 291
Db 200 ATFAVRSGLLTVMWVSSVFFFPVFCULTVLSLIGRKLW---RRKRGNAVGSLSRDQ 256

Qy 292 GHRQTVRVLKWRSGSKDACLQAPPG 319

```

Db 257 NHRQTVKML-----GGSQCALELSLPG 278

RESULT 12

US-10-303-204A-16

; Sequence 16, Application US/10303204A

; Publication No. US20030166144A1

; GENERAL INFORMATION:

; APPLICANT: Arena, Joseph P.

; APPLICANT: Cully, Doris F.

; APPLICANT: Feighner, Scott D.

; APPLICANT: Howard, Andrew D.

; APPLICANT: Liberator, Paul A.

; APPLICANT: Schaeffer, James M.

; APPLICANT: Van Der Ploeg, Leonardus H. T.

; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR

; TITLE OF INVENTION: FAMILY

; FILE REFERENCE: 19589PCA

; CURRENT APPLICATION NUMBER: US/10/303,204A

; CURRENT FILING DATE: 2002-11-25

; PRIOR APPLICATION NUMBER: 09/077,674

; PRIOR FILING DATE: 1998-06-03

; PRIOR APPLICATION NUMBER: PCT/US96/19445

; PRIOR FILING DATE: 1996-12-10

; PRIOR APPLICATION NUMBER: 60/018,962

; PRIOR FILING DATE: 1996-06-06

; PRIOR APPLICATION NUMBER: 60/008,582

; PRIOR FILING DATE: 1995-12-13

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 16

; LENGTH: 364

; TYPE: PRT

; ORGANISM: rattus norvegicus

US-10-303-204A-16

Query Match 31.4%; Score 641; DB 12; Length 364;

Best Local Similarity 45.8%; Pred. No. 3e-43;

Matches 142; Conservative 41; Mismatches 65; Indels 62; Gaps 10;

QY 5 WNGSDGPEGAREP-----WPALPPCD-----ERRCSPFPLGALVPVAVCLCLFVVGVGSG 55

Db 2 WNAT--PSEPEPENVTLDDWDASPGNDSLPLDELLPLFPAPLAGVTATCVALFVVGISG 59

QY 56 NVVTMLIGRYDRMTTNLYLGSMAVSDLLILGLPFDLYRLWRSRPWVFGPLCLSL 115

Db 60 NLUTMLVSVRFRELTNTNLYLGSMAFSDLLILFCLMPLDLVLRWQYRPNWVAFCSAGPIFV 119

QY 116 YVGEQCTYATLLHMTALSVERYLAI CRPLRARVLTERRVRALIAVLWAVALLSAGPFLF 175

Db 120 FVSECTYATVLTITALSVERYLAI CRPLRARVLTERRVRALIAVLWAVALLSAGPFLF 179

QY 176 LVGVEQDPGISVVPGLNGTARIASSPLASSPPLWLRAPPPSPGPETAALFSPREC 235

Db 180 LVGVEHE-----NGT-----DPRD-----TNEC 197

QY 236 RPS--PAQGLARVLMWVTAYTFPLFPLCLSLYGLIGRELWSSRRPLRGPAASG---RE 290

Db 198 RATEFAVRSGLLTVMWVSSVFFFLPVCLTVLYSLIGRKLW--RR--RGDAAVGASLRD 253

QY 291 RGHRTQTVRVL 300

Db 254 QNHQTVKML 263

RESULT 13

US-10-303-204A-3

; Sequence 3, Application US/10303204A

; Publication No. US20030166144A1

; GENERAL INFORMATION:

; APPLICANT: Arena, Joseph P.

; APPLICANT: Cully, Doris F.

; APPLICANT: Feighner, Scott D.

; APPLICANT: Howard, Andrew D.

; APPLICANT: Liberator, Paul A.

; APPLICANT: Schaeffer, James M.

; APPLICANT: Van Der Ploeg, Leonardus H. T.

; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR

; TITLE OF INVENTION: FAMILY

; FILE REFERENCE: 19589PCA

; CURRENT APPLICATION NUMBER: US/10/303,204A

; CURRENT FILING DATE: 2002-11-25

; PRIOR APPLICATION NUMBER: 09/077,674

; PRIOR FILING DATE: 1998-06-03

; PRIOR APPLICATION NUMBER: PCT/US96/19445

; PRIOR FILING DATE: 1996-12-10

; PRIOR APPLICATION NUMBER: 60/018,962

; PRIOR FILING DATE: 1996-06-06

; PRIOR APPLICATION NUMBER: 60/008,582

; PRIOR FILING DATE: 1995-12-13

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3

; LENGTH: 353

; TYPE: PRT

; ORGANISM: sus scrofa

US-10-303-204A-3

Query Match 31.3%; Score 639.5; DB 12; Length 353;

Best Local Similarity 47.2%; Pred. No. 3.8e-43;

Matches 137; Conservative 39; Mismatches 61; Indels 53; Gaps 7;

QY 19 WPALPPCD---ERRCSPFPLGALVPVAVCLCLFVVGVSGNVVTVMLIGRYDRMTTNL 75

Db 8 WDAPPENDSLVEELLPLFPPLLAGVTATCVALFVVGIAAGNLLMLVSRFRMTTNL 67

QY 76 YLGSMVSDLLILGLPFDLYRLWRSRPWVFGPPLCLSLYVGEQCTYATLLHMTALSVE 135

Db 68 YLSSMAFSDLLILFCLMPLDLVLRWQYRPNWVAFCSAGPIFVVGVEHD 127

QY 136 RYLAI CRPLRARVLTERRVRALIAVLWAVALLSAGPFLFVGVQDPGISVVPGLNGTA 195

Db 128 RFALCFPLRAKVVTGKRVILVWVAFCSAGPIFVVGVEHD-----NGT- 177

QY 196 RIASSPLASSPPLWLRAPPPSPGPETAALFSPRECPS--PAQGLARVLMWVT 253

Db 178 -----DPRD-----TNECATEFAVRSGLLTVMWVSS 205

QY 254 AVFFLPCLSLYGLIGRELWSSRRPLRGPAASG---RERGHRTQTVRVL 300

Db 206 VFFFLPVCLTVLYSLIGRKLW---RRKRGAAVGSRLRDQNHQTVKML 252

RESULT 14

US-10-303-204A-8

; Sequence 8, Application US/10303204A

; Publication No. US20030166144A1

; GENERAL INFORMATION:

; APPLICANT: Arena, Joseph P.

; APPLICANT: Cully, Doris F.

; APPLICANT: Feighner, Scott D.

; APPLICANT: Howard, Andrew D.

; APPLICANT: Liberator, Paul A.

; APPLICANT: Schaeffer, James M.

; APPLICANT: Van Der Ploeg, Leonardus H. T.

; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR

; TITLE OF INVENTION: FAMILY

; FILE REFERENCE: 19589PCA

; CURRENT APPLICATION NUMBER: US/10/303,204A

; CURRENT FILING DATE: 2002-11-25

; PRIOR APPLICATION NUMBER: 09/077,674

; PRIOR FILING DATE: 1998-06-03

; PRIOR APPLICATION NUMBER: PCT/US96/19445

; PRIOR FILING DATE: 1996-12-10

; PRIOR APPLICATION NUMBER: 60/018,962

; PRIOR FILING DATE: 1996-06-06

PRIOR APPLICATION NUMBER: 60/008,582
; PRIOR FILING DATE: 1995-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 361
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-303-204A-8

Query Match 31.2%; Score 639.5; DB 12; Length 361;
Best Local Similarity 47.0%; Pred. No. 3.9e-43;
Matches 135; Conservative 41; Mismatches 64; Indels 47; Gaps 6;

Db	Seq	Score	Length	Indels	Gaps
QY	19 WPALPPCD---ERRCSPPFGALVPTAVCLCLFVVGVSGNVVTVMLIGRYDRMTTNNL	75			
Db	16 WDASPGNDSLGDELLQLFPAPLAGVTATCVAFVVGIAGNLLTMLVVSFRRLRTTNNL	75			
QY	76 YLGSMAVSDLLIILGLPFDLYRLWRSRPWPVFGPGLLCRLSLYVGEGCTYATLLHMTALSVE	135			
Db	76 YLSSMAFSDLLIFLCMPDLVRLWQYRPMNFGDLCCKLFQFVSESCCTYATVLTITALSVE	135			
QY	136 RYLAICRPLRARBVLTRRRVRLIALVLMVALLSAGPFLVGVQDPGIVVVGPNLNGTA	195			
Db	136 RYFAICPPLRAKVVVTKGRVKLVIFVIWAFCSAGPIFVLVGVGHE-----NGT- 185				
QY	196 RIASSPLASSPPLWLSRAPPPSPGPETAEEAALFSRECRPS--PAQLGALRMLWVTT	253			
Db	186 -----DP--W-----DINECRPTEFAVRSGLLTVVWVSS 213				
QY	254 AYFPLPCLSLYGLIGRLWSRRRLRGPAAASGRGRGHRQTVRVL 300				
Db	214 IFFFLPVFCLTVLSYLGRLKLRRRRGDAVVGASLRDQNHKQTVKML 260				

RESULT 15
US-10-251-385-210
; Sequence 210, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
; TITLE OF INVENTION: Protein-Coupled
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/10/251,385
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US/09/170,496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 210
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-251-385-210

Query Match 31.2%; Score 637; DB 15; Length 366;
Best Local Similarity 44.8%; Pred. No. 6.2e-43;
Matches 137; Conservative 43; Mismatches 74; Indels 52; Gaps 7;

Db	Seq	Score	Length	Indels	Gaps
QY	5 WNGSDGPEGA-----REPPWALPPCD---ERRCSPPFGALVPTAVCLCLFVVGVSGN	56			
Db	2 WNATPSEPGFNLTADLDWDASGNDLSGDELLQLFPAPLAGVTATCVAFVVGIAGN	61			
QY	57 VVTVMILIGRYDRMTTNNLYLGSMAVSDLLIILGLPFDLYRLWRSRPWPVFGPGLLCRLSLY	116			
Db	62 LLTMLVVSFRRLRTTNNLYLSSMAFSDLLIFLCMPDLVRLWQYRPMNFGDLCCKLFQF	121			
QY	117 VBGECCTYATLLHMTALSVERYLAICRPLRARBVLTRRRVRLIALVLMVALLSAGPFLFL	176			

Search completed: January 1, 2004, 06:53:29
Job time : 142.727 sec

Search completed: January 1, 2004, 06:53:29
Job time : 142.727 sec

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 1, 2004, 06:20:00 ; Search time 24.6692 Seconds
(without alignments)
1504.757 Million cell updates/sec

Title: US-09-719-485-5
Perfect score: 2043
Sequence: 1 MGSPWNGSDGEGAREPPWP.....MQNHLKHGRFADVLLSVL 386

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	343	16.8	424	2 JH0164	neurotensin recept
2	330	16.2	418	2 S29506	neurotensin recept
3	316	15.5	477	2 JC7913	capa receptor (CGI
4	285.5	14.0	418	2 A88013	protein K10B4.4 [1
5	268.5	13.1	416	2 S68822	neurotensin recept
6	255.5	12.5	378	2 T15816	hypothetical prote
7	253.5	12.4	367	2 I49022	kappa opioid recep
8	253.5	12.4	367	2 JC2421	opioid receptor ho
9	253.5	12.4	367	2 I56520	G protein-coupled
10	253.5	12.4	370	2 S43087	orphan opioid rece
11	250.5	12.3	363	2 I57940	somatostatin recep
12	249.5	12.2	452	2 A46195	cholecystokinin B
13	249.5	12.2	452	2 JC2459	gastrin/cholecysto
14	248.5	12.2	418	2 A46226	somatostatin recep
15	247	12.1	453	2 S32817	gastrin receptor -
16	245.5	12.0	363	2 I57955	somatostatin recep
17	245.5	12.0	364	2 JN0763	somatostatin recep
18	243	11.9	359	2 JC5277	G protein-coupled
19	241.5	11.8	477	1 QRHUB1	beta-1-adrenergic
20	240	11.7	450	2 JQ1614	gastrin receptor -
21	239	11.7	352	2 JE0296	thyrotropin releas
22	237.5	11.6	447	2 A47430	gastrin/cholecysto
23	237.5	11.6	519	2 S17783	tachykinin recepto
24	234.5	11.5	514	2 D56849	dopamine receptor-
25	233.5	11.4	480	2 I53053	beta 1 adrenergic
26	232.5	11.4	379	2 JC6178	serotonin receptor
27	232.5	11.4	450	2 I49481	alpha-2 adrenergic
28	232	11.4	380	2 A55259	kappa opioid recep
29	230	11.3	477	2 T25846	hypothetical prote

30	229.5	11.2	466	2 S36794	beta-1-adrenergic
31	229	11.2	427	2 S50150	gastric CCK-A rece
32	228.5	11.2	387	2 JC5949	galanin receptor 2
33	228	11.2	437	2 I57942	5-hydroxytryptamin
34	228	11.2	450	2 B40392	alpha-2-adrenergic
35	227.5	11.1	428	2 S30508	probable G protein
36	227	11.1	380	2 JC2338	kappa opioid recep
37	227	11.1	418	2 G02953	beta-3-adrenergic
38	226	11.1	398	2 JN0708	thyrotropin-releas
39	225	11.0	444	2 A42685	cholecystokinin re
40	224.5	11.0	411	2 I56444	thyrotrophin-relea
41	224.5	11.0	412	2 S23436	thyroliberin recep
42	224.5	11.0	428	2 A44021	somatostatin recep
43	224	11.0	483	2 A25896	beta-adrenergic re
44	223.5	10.9	436	2 JC5599	cholecystokinin-A
45	223.5	10.9	450	2 A38316	alpha-2-adrenergic

ALIGNMENTS

RESULT 1

JH0164

neurotensin receptor - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 17-Mar-2000

C;Accession: JH0164

R;Tanaka, K.; Masu, M.; Nakanishi, S.

Neuron 4, 847-854, 1990

A;Title: Structure and functional expression of the cloned rat neurotensin receptor.

A;Reference number: JH0164; MUID:90297956; PMID:1694443

A;Accession: JH0164

A;Molecule type: mRNA

A;Residues: 1-424 <TAN>

C;Comment: Neurotensin receptor belongs to the family of G protein-coupled receptor. This

ter (neuromodulator) in the brain and as a hormone) cellular mediator in peripheral tiss

C;Superfamily: vertebrate rhodopsin

C;Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein

F;65-87/Domain: transmembrane #status predicted <TM1>

F;97-121/Domain: transmembrane #status predicted <TM2>

F;144-165/Domain: transmembrane #status predicted <TM3>

F;189-210/Domain: transmembrane #status predicted <TM4>

F;236-260/Domain: transmembrane #status predicted <TM5>

F;309-330/Domain: transmembrane #status predicted <TM6>

F;348-372/Domain: transmembrane #status predicted <TM7>

F;4,38,42/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 16.8%; Score 343; DB 2; Length 424;
Best Local Similarity 33.9%; Pred. No. 4.7e-19;
Matches 82; Conservative 41; Mismatches 71; Indels 48; Gaps 6;

QY	39	VPVTAVCLCLFVVGSGNVVTVMLIGR---YRDMRTTNLYLGSMVSDLLI-LLGLPPD 94
DB	65	VLVTALYALFVVGTVGNSVTAFTLARKSLQSLQSTVHYHLGSLALSLLILLAMPVE 124
QY	95	LYR-LWRSPWVFGPLLCRLSLVYVGSGCTVATLHMTALSVRYLAICRPLRVLVTRR 153
DB	125	LYNFIWHPWAGDAGCGYFRLDACTYATALNVASLSVERYLAICHPKATLMRSR 184
QY	154	RVRALIAVLMAVALLSAGPFLFLVVGVEQDPGIVSVFGLNGTARIASSPLMLSRA 213
DB	185	RTKKFISAIWLASALLAIPMLFTMLQNRSGDGTHPG-----GLVCTPIVDT----- 231
QY	214	PPSPSGPETAATAAALFSECRPSPAQLGALRVMLVWTYATYFPL-PPCLSLIYLGLIGR 272
DB	232	-----ATVKVIVQNTFMGFLPMLVISILNTVIAN 262
QY	273	EL 274
DB	263	KL 264

RESULT 2

A;Accession: S69822
A;Molecule type: mRNA
A;Residues: 1-416 <CHA>
A;Cross-references: GB:X97121; NID:g1483579; PIDN:CAA65787.1; PID:g1483580
A;Experimental source: hypothalamus
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; transmembrane protein
F;33-59/Domain: transmembrane #status predicted <TM1>
F;70-91/Domain: transmembrane #status predicted <TM2>
F;110-131/Domain: transmembrane #status predicted <TM3>
F;155-175/Domain: transmembrane #status predicted <TM4>
F;204-230/Domain: transmembrane #status predicted <TM5>
F;296-315/Domain: transmembrane #status predicted <TM6>
F;335-361/Domain: transmembrane #status predicted <TM7>

Query Match 13.1%; Score 268.5; DB 2; Length 416;
Best Local Similarity 37.9%; Pred. No. 2.6e-13;
Matches 66; Conservative 33; Mismatches 66; Indels 9; Gaps 4;

QY 18 PWPALP-----PCDERCSPFPLGALVPVAVCLCLFVVGVSGNVVTVMLIGRYDRMT 71
Db 6 PWPSPSPSAGLSLEARLGVDTLWAKVLTALYSILFAFGTAGNALSVMHVKARAGRP 65

QY 72 -TTNLYLGSMAVSDLLILL-GLPFDLYR-LWRSRPWVFGPLCLSLYVGEGETYATLH 128
Db 66 GLRYHVLVSALGALLLLVSMPELNFVMSHPWVFGDLGCRGYFVRELCAATVLS 125

QY 129 MTALSVERYLAICRPLRARVLTTRRRVRLIALIWAVALLSAGPFLFLVGVQD 182
Db 126 VASLSAERCLAVCQPLRARLLTPRRTRRLSLVWVASLGLALPMAVIMQKHE 179

RESULT 6
T15816
hypothetical protein C48C5.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Mar-2000
C;Accession: T15816
R;Favella, A.
submitted to the EMBL Data Library, November 1995
A;Description: The sequence of C. elegans cosmid C48C5.
A;Reference number: Z18410
A;Accession: T15816
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-378 <FAV>
A;Cross-references: EMBL:U39994; NID:g1055102; PID:g1055105; PIDN:AAB37017.1; GSPDB:GN00
A;Experimental source: strain Bristol N2; clone C48C5
C;Genetics:
A;Gene: CESP:C48C5.1
A;Map position: X
A;Introns: 53/3; 87/2; 136/2; 169/3; 209/3; 231/3; 259/3; 286/1; 327/1
C;Superfamily: adenosine receptor A1

Query Match 12.5%; Score 255.5; DB 2; Length 378;
Best Local Similarity 27.0%; Pred. No. 2.4e-12;
Matches 66; Conservative 49; Mismatches 96; Indels 33; Gaps 6;

QY 38 LVPVAVCLCLFVVGVSGNVVTVMLIGRYDRMTTNNLYLGSMAVSDLLIL-LGLPDLY 96
Db 51 LYKVATLYIFLFGVIGNTTCLVMKKHPMMKTHASMLNLAVALSDVLTLCVGLPEVM 110

QY 97 RLWRSRPWVFGPLCLSLYVGEGETYATLHMTALSVERYLAICRPL-RARVLTTRRV 155
Db 111 MNWQYWPFPDYICNLKALIAETSSVSLITLIFAIERYVAVCHFLMKVQPPKRN 170

QY 156 RLIALIWAVALLSAGPFLFLVGVQDPCISVVPGLNGTARIASSPPLWLSRAP 215
Db 171 GTIIGTWFISILCAMP--FAIHRADYIMKSWGTDNRIPVKSCKM----- 217

QY 216 PSPSPGPETAALFSPRCRPSPAQLGALRVML-WYTTAYFFLPFLCLSLYGLIGREL 274
Db 218 -----IAVWF-----EPKLASTKILFHSALIAFFALPLFTIVILYARIACKV 260

QY 275 WSSR 278
Db 261 SSNR 264

RESULT 7
I49022
kappa opioid receptor 3 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 19-May-2000
C;Accession: I49022
R;Pan, Y.X.; Cheng, J.; Xu, J.; Rossi, G.; Jacobson, E.; Ryan-Moro, J.; Brookes, A.I.; De
Mol. Pharmacol. 47, 1180-1188, 1995
A;Title: Cloning and functional characterization through antisense mapping of a kappa 3-
A;Reference number: I49022; MUID:95327076; PMID:7603458
A;Accession: I49022
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: mRNA
A;Residues: 1-367 <RES>
A;Cross-references: EMBL:U09421; NID:g551484; PIDN:AAA81333.1; PID:g551485
C;Superfamily: vertebrate rhodopsin

Query Match 12.4%; Score 253.5; DB 2; Length 367;
Best Local Similarity 26.8%; Pred. No. 3.3e-12;
Matches 72; Conservative 47; Mismatches 95; Indels 57; Gaps 6;

QY 34 PLGALVPVAVCLCLFVVGVSGNVVTVMLIGRYDRMTTNNLYLGSMAVSDLLILCLPFP 93
Db 44 PLGLKVTIVGLYLVAVCIGLLGCLVMYILRHTKMTATNIYIFNLALADTLVLTLPF 103

QY 94 ---DLYRLWRSRPWVFGPLCLSLYVGEGETYATLHMTALSVERYLAICRPLRARV 150
Db 104 QGTDILLGP-----WPFGNALCKTVIAIDYVNMFTSTFTLTAMSDRVVAICHPDALDVR 159

QY 151 TRRVVATLAVWALLSAGPFLFLVGVQDPCISVVPGLNGTARIASSPPLW 210
Db 160 TSSKAQAVNVAIWALASV-----VGVPVAIMGSAQVDEEIEC-----L 198

QY 211 SRAPPPSPSPGPTAAALFSPRCRPSPAQLGALRVMLWYTTAYFPFLCLSLYGLI 270
Db 199 VEIPAPQDYWGVPVFAICIFLFS-----FIIPVLIIISVCYSML 235

QY 271 GRELWSSRRPLRGPAASGRGRGHRTQVRVLR 301
Db 236 IRLRGVR-----LLSGSRKDRNLRRT 260

RESULT 8
JC2421
opioid receptor homolog, MOR-C - mouse
C;Species: Mus musculus (house mouse)
C;Date: 21-Feb-1995 #sequence_revision 05-Apr-1995 #text_change 24-Nov-1999
C;Accession: JC2421; I49122
R;Nishi, M.; Takeshima, H.; Mori, M.; Nakagawara, K.; Takeuchi, T.
Biochem. Biophys. Res. Commun. 205, 1353-1357, 1994
A;Title: Structure and chromosomal mapping of genes for the mouse kappa-opioid receptor
A;Reference number: JC2421; MUID:95100967; PMID:7802669
A;Accession: JC2421
A;Molecule type: mRNA
A;Residues: 1-367 <NIS>
A;Cross-references: DDBJ:D31663
R;Halford, W.P.; Gebhardt, B.M.; Carr, D.J.
J. Neuroimmunol. 59, 91-101, 1995
A;Title: Functional role and sequence analysis of a lymphocyte orphan opioid receptor.
A;Reference number: I49122; MUID:95318231; PMID:7797625
A;Accession: I49122
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: mRNA
A;Residues: 1-357 <RES>
A;Cross-references: EMBL:U14165; NID:g540092; PIDN:AAA87899.1; PID:g540093
C;Genetics:
A;Gene: MOR-c; OOR

C;Keywords: G protein-coupled receptor; transmembrane protein

Query Match 12.4%; Score 253.5; DB 2; Length 370;
 Best Local Similarity 26.9%; Pred. No. 3.3e-12;
 Matches 73; Conservative 46; Mismatches 95; Indels 57; Gaps 6;
 QY 34 PLGALVPVAVCLCLFVGVSGNNVTMLIGRYDRMRTNTNLYLGSMAVSDLLILGLLP 93
 DB 47 PLGLKVTIVGLYLAVCVGLGNCVYVILTRKMTATNIYFNALADTLVLTLP 106
 QY 94 ---DLVLRWSRPVSGPLLCRLSLVYVGGCTATLLHMTALSVRYLAICRPLRVLV 150
 DB 107 QGTDIILGP-----WPGNALCKTVIAIDYNNMTSTFTLTMVSDRYVAICHPRALDVR 162
 QY 151 TRRRVRALIAVMAVALLSAGPFLFLVGVQDPGCVSVVPGNLGTARIASSPLASSPPLWL 210
 DB 163 TSSKAQAVNVAIWALASV-----VGVPVAIMGSAQVDEIEEC-----L 201
 QY 211 SRAPPPSPGPTABAAALFSRECRPSPAQLGALRVMLVWTYATPFLPCLSLIYGLI 270
 DB 202 VEIPTQDYMGVPFAICIFLFS-----FIVPVLVISVCYSLM 238
 QY 271 GRELWSRRRLRCPAASGRGRHQTVRVLR 301
 DB 239 IRLRGVR-----LUSGSRKDRNLRRITR 263
 RESULT 11
 157940
 somatostatin receptor 5 - rat
 N;Alternate names: somatotropin release-inhibiting factor subtype 28 receptor
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 02-Aug-1996 #sequence revision 02-Aug-1996 #text_change 24-Nov-1999
 C;Accession: I57940; S39244
 R;O'Carroll, A.M.; Lolait, S.J.; Konig, M.; Mahan, L.C.
 Mol. Pharmacol. 42, 939-946, 1992
 A;Title: Molecular cloning and expression of a pituitary somatostatin receptor with pref
 A;Reference number: I57940; MUID:93125499; PMID:1362243
 A;Accession: I57940
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-363 <CAL>
 A;Cross-references: GB:I04535; NID:G409238; PIDN:AAAL7029.1; PID:G409239
 R;O'Carroll, A.M.; Lolait, S.J.; Konig, M.; Mahan, L.C.
 Mol. Pharmacol. 44, 1278, 1993
 A;Title: Molecular cloning and expression of a pituitary somatostatin receptor with pref
 A;Reference number: I57949; MUID:94088493; PMID:8264565
 A;Accession: I57949
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 341-363 <CA2>
 A;Cross-references: GB:S67370; NID:G455947; PIDN:AAB29371.1; PID:G455948
 A;Experimental source: pituitary
 R;Fenicta, R.; Greenwood, M.; Patel, Y.C.
 Submitted to the EMBL Data Library, August 1993
 A;Description: Correction of the nucleotide and amino acid sequence of the rat somatosta
 A;Reference number: S39244
 A;Accession: S39244
 A;Molecule type: mRNA
 A;Residues: 309-363 <PEN>
 A;Cross-references: EMBL:X74828; NID:G433911; PIDN:CAA52825.1; PID:G433912
 C;Genetics:
 A;Gene: SSTR5
 C;Superfamily: vertebrate rhodopsin

Query Match 12.3%; Score 250.5; DB 2; Length 363;
 Best Local Similarity 26.7%; Pred. No. 5.5e-12;
 Matches 76; Conservative 49; Mismatches 99; Indels 67; Gaps 8;
 QY 5 WNSDGPGEAREPWPALPCDERRCSPFLGA---LVPVTAIVCLCLFVGVSGNNVTVM 61
 DB 12 WNASASSGNHN---WSLVG-----SASPMGARAVLPVLYLLVC---TVGLSGNTLIVY 60

QY 62 LIQRYDRMRTNTNLYLGSMAVSDLLILGLPDLVLRWSRPVWPGPLLCRLSLYVGECC 121
 DB 61 VVLRHAKMTVTNVTINLAVADVLFMGLPFLATQNAVVSYPFGSFLCLRLVMTLDGIN 120
 QY 122 TYATLLHMTALSVRYLAICRPLRVLTTRRRVALIAVMAVALLSAGPFLFLVGVQ 181
 DB 121 QFTSIPCLMVMVSDRYLAIVHPLRSARWRPRVAKMASAAVWVFSLSMELPLLVPAVQ 180
 QY 182 DRGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAEAAALFSRECRPSPAQ 241
 DB 181 G-----WGTCNLS-----W-----PEPVG 194
 QY 242 LGALRVMLVWTYATPFLPCLSLIYGLI-----GRELWSRR 279
 DB 195 LMGAAPITTYTSLVGLPFLVLCCLYLLIIVKVKAGMRVGSRR 239
 RESULT 12
 A46195
 cholecystokinin B receptor subtype - rat
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 21-Sep-1993 #sequence revision 18-Nov-1994 #text_change 20-Apr-2000
 C;Accession: A46195
 R;Wank, S.A.; Pisegna, J.R.; de Weerth, A.
 Proc. Natl. Acad. Sci. U.S.A. 89, 8691-8695, 1992
 A;Title: Brain and gastrointestinal cholecystokinin receptor family: structure and functi
 A;Reference number: A46195; MUID:92409582; PMID:1528881
 A;Accession: A46195
 A;Status: preliminary
 A;Molecule type: nucleic acid
 A;Residues: 1-452 <WAN>
 A;Cross-references: GB:M99418; NID:G203459; PIDN:AAA40925.1; PID:G203460
 A;Experimental source: brain
 A;Note: sequence extracted from NCBI backbone (NCBIN:114083, NCBI:114084)
 C;Superfamily: neurokinin 1 receptor
 C;Keywords: G protein-coupled receptor; transmembrane protein
 Query Match 12.2%; Score 249.5; DB 2; Length 452;
 Best Local Similarity 26.1%; Pred. No. 8.3e-12;
 Matches 84; Conservative 51; Mismatches 96; Indels 91; Gaps 13;
 QY 48 LFWGVSGNNVTMLIGRYDRMRTNTNLYLGSMAVSDLLILGL-LPPDLVLRWSRP--- 103
 DB 64 IFLMSVGGNLIIVLVGLSLRRLTNTAFNLSLAVSDLLAVACMPFTLL-----PNLM 117
 QY 104 --WVFGPLLCR-LSLVYVGGCTATLLHMTALSVRYLAICRPLRVLTTRRRVALIA 160
 DB 118 GTFIFGTIVCKAISYLMGVSVSST-LNLVAIALERYSAICRPLQARVWQTRSHAARVIL 176
 QY 161 VLMAVALLSAGPFLFLVGVQDPGCVSVVPGNLGTARIASSPLASSPPLWLSRAPPPSPPS 220
 DB 177 ATWLLSGLLMVPI-----PVYTMV-----QPV 198
 QY 221 GPETAABAAALFSRECRPSPAQLGALRVMLVWTYATPFLPCLSLIYGLIGRELW----- 275
 DB 199 GPRVLCQMRW-----PSARVQQTQSVLILL--LLPFIQGVIAVAYGLSRELYLGLHF 251
 QY 276 -----SSRRPLGPAASG---RERGRHQTVRVLRKWSRSGSKDACLQSPAPGT 320
 DB 252 DGENDETOSRARNQGLPGGAAPGVHONGGCRPVTSVAGE-----DSGCCVQLPRSR 306
 QY 321 AQTGLPLPQLLAQLWAPLPAPFP 342
 DB 307 LE-----MTLTITPTPGVP 321
 RESULT 13
 JC2459
 gastrin/cholecystokinin B receptor - rabbit
 C;Species: Oryctolagus cuniculus (domestic rabbit)
 C;Date: 21-Feb-1995 #sequence revision 05-Apr-1995 #text_change 20-Apr-2000
 C;Accession: JC2459
 R;Blandizzi, C.; Song, I.; Yamada, T.

Biochem. Biophys. Res. Commun. 202, 947-953, 1994

A;Title: Molecular cloning and structural analysis of the rabbit gastrin/CKK receptor gene
A;Reference number: J02459; MUID:94324990; PMID:8048969
A;Accession: J02459
A:Molecule type: mRNA
A;Residues: 1-452 <BLA>
A;Cross-references: GB:IJL1548; NID:g495663; PIDN:AAA31194.1; PID:g495665
C;Genetics:
A;Introns: 49/1; 133/1; 216/2; 273/1
C;Superfamily: neurokinin 1 receptor
C;Keywords: receptor; transmembrane protein
F;56-79/Domain: transmembrane #status predicted <TM1>
F;85-104/Domain: transmembrane #status predicted <TM2>
F;130-149/Domain: transmembrane #status predicted <TM3>
F;169-187/Domain: transmembrane #status predicted <TM4>
F;217-237/Domain: transmembrane #status predicted <TM5>
F;339-359/Domain: transmembrane #status predicted <TM6>
F;381-400/Domain: transmembrane #status predicted <TM7>

Query Match 12.2%; Score 249.5; DB 2; Length 452;
Best Local Similarity 25.2%; Pred.No. 8.3e-12;
Matches 86; Conservative 50; Mismatches 122; Indels 83; Gaps 10;

QY 48 LFVGVGSGNVVTVMILIGRYEDMTTNNLYLGSAVSLLLLILGL-LPDLIYRLWRSRP--- 103
DB ::: :::: :||| :||| :||| :||| :||| :||| :||| :||| :|||
62 IFMSVGNGNTLLIIIVVLGSRLRTVTNAFLSLAVSDLLLAACMPFTLL-----PNLM 115
QY 104 --WVFPGLLCLRLSLVYGEGCTYATLLHMTALSVERYLAI CRPLRARVLVTRRRVALIAV 161
DB 116 GTIFGTGVICAKSYLMGVSVSVSTLSLVAILERTYSACRPLQARVWQTRSHAARVILA 175
QY 162 LWAVALLSAGPFLLVLGVGEODPGISVVVPLGNGTARIASSPLASSPPMLWSRAPPPSPSG 221
DB 176 TLLSLGLMVPYPVTVTAQ-----PVG 197
QY 222 PETAEAAALFSRECRPPSPAQLGALRVMLWTYYFFLPFLCLSYLCIGRELWSSRR-- 279
DB 198 PRVLQCVRHW-----PSARVQTWTSVLLLRL--LLFFVPGVVMAYAVGLISRLEYLGLRFD 250
QY 280 -----PLRGPAASGRERGHQTVRVLRKWSRRGSK-DACLQSAPPGTQAOT 323
DB 251 SDSDESQSQRVGGGGLPGGAAGFPV---HQNGRCRPEAGLEDGDGCYVOLPRSR--- 304
QY 324 LGPLPLIAQLWAPLPAPFPPTISI PASTRRGGSGGYINLLVAL 364
DB 305 --PALELSALTAPTISGPGPRPAQAOKLLAKKRVRMILLVI 343

RESULT 14
A46226
somatostatin receptor 3 - human
C;Species: Homo sapiens (man)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 24-Nov-1999
C;Accession: A46226; S32501
R;Yamada, Y.; Reisine, T.; Law, S.F.; Ihara, Y.; Kubota, A.; Kagimoto, S.; Seino, M.; Seino, M.; Endocrinol. 6, 2136-2142, 1992
A;Title: Somatostatin receptors, an expanding gene family: cloning and functional characterization
A;Reference number: A46226; MUID:93149123; PMID:1337145
A;Accession: A46226
A;Molecule type: DNA
A;Residues: 1-418 <YAM>
A;Cross-references: GB:M96738; NID:G338498; PID:G338499
A;Note: sequence extracted from NCBI backbone (NCBIN:123685, NCHIP:123690)
R;Corness, J.D.; Demchyshyn, L.L.; Seeman, P.; van Tol, H.H.M.; Srikant, C.B.; Kent, G.; FEBS Lett. 321, 279-284, 1993
A;Title: A human somatostatin receptor (SSTR3), located on chromosome 22, displays preferential binding to the [D-Trp6]somatostatin-14
A;Reference number: S32501; MUID:93238970; PMID:8097479
A;Accession: S32501
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-418 <COR>
C;Genetics:
A;Gene: GDB:SSTR3

A;Cross-references: GDB:I34187; OMIM:182453
A;Map position: 22q13.1-22q13.1
A;Introns: #status absent
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane protein
F;44-70/Domain: transmembrane #status predicted <TM1>
F;81-106/Domain: transmembrane #status predicted <TM2>
F;118-139/Domain: transmembrane #status predicted <TM3>
F;159-181/Domain: transmembrane #status predicted <TM4>
F;203-233/Domain: transmembrane #status predicted <TM5>
F;255-282/Domain: transmembrane #status predicted <TM6>
F;289-316/Domain: transmembrane #status predicted <TM7>
F;17,30/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;116-191/disulfide bonds: #status predicted
F;151,251,317,332/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted
F;251/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #status predicted
F;256/Binding site: phosphate (Thr) (covalent) (by cAMP-dependent kinase) #status predicted
F;412/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted

Query Match 12.2%; Score 248.5; DB 2; Length 418;
Best Local Similarity 23.7%; Pred. No. 9.1e-12;
Matches 88; Conservative 51; Mismatches 133; Indels 99; Gaps 11;

Qy	11	PEGAREPPWALPPCDERRCSFPPL-----GALVPVTAVCLCLFVVGSGNVVTVMLIGR	65
Db	15	PENA--SSAWPDPATGNVSAGSPAGLAVSGVLIPLVYLCVC--VVGLLGNSLIYYVCLR	71
Qy	66	YRD MRTTTNL YLGSMASVDLLILGLLPDLYRLMRSRPWFVGPLLCRLSLYVGGEGCTYAT	125
Db	72	HTAPSVTNVYIINLALADELFPMGLPF--LAAQNALSYPFGSLMCRLVMAVDGINQFTS	130
Qy	126	LHNTALSYERYLAICRPLRARLVTRRRVRALIIVLWALLSAGFFFLVGVGEQDGI	185
Db	131	IFCLTMSVDRYLVAVHPTRSAWRPTAPVARTVSAAVWVASAVVLPVVVFSGV-----	184
Qy	186	SWPFLNGTARIASSPLASSPPLWLSRAPPPPGSPGPETAFAAAAFSRECK---PSPAQL	242
Db	185	-----PRGMST-----CHMQWPETAAA	201
Qy	243	GALRVMLWTYAYFPFLPFLCLISLYGLIGRELWSSRRPLRGPAASGRGERGHQTIVLRK	302
Db	202	WRAGFIITTAALGFPGFLPVICLCYLIIIVKVSAGRVRWAPSCORRRRSERRVTRMV--	259
Qy	303	WSRRGSKDACIQSAPPGTAGTLGPLLALWAP-----LPAPPISIPASTRRGGG	354
Db	260	-----VAVVFLVLCWMPFYVLNINVVCCPLP-EEPAPF-----	291
Qy	355	SGIYNLLVALP	365
Db	292	FGLYFLVALP	302

RESULT 15
S32817
gastrin receptor - dog
C;Species: Canis lupus familiaris (dog)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Apr-2000
C;Accession: S32817
R;Kopin, A.S.; Lee, Y.M.; McBride, E.W.; Miller, L.J.; Lu, M.; Lin, H.Y.; Kolakowski Jr.
Proc. Natl. Acad. Sci. U.S.A. 89, 3605-3609, 1992
A;Title: Expression cloning and characterization of the canine parietal cell gastrin recep
A;Reference number: S32817; PMID:92228835; PMID:1373504

A;Residues: 1-453 <KOP>
A;Cross-references: EMBL:M87834; NID:gl63956; PIDN:AAA30847.1; PID:gl63957
C;Superfamily: neurokinin 1 receptor
C;Keywords: G protein-coupled receptor; transmembrane protein
Query Match 12.1%; Score 247; DB 2; Length 453;
Best Local Similarity 25.9%; Pred. No. 1.3e-11;
Matches 97; Conservative 50; Mismatches 110; Indels 118; Gaps 13;

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 1, 2004, 06:19:59 ; Search time 13.5439 Seconds
(without alignments)
1340.261 Million cell updates/sec

Title: US-09-719-485-5
Perfect score: 2043
Sequence: 1 MGSPWNGSDGPEGAREPPWP.....WQNHKHKGRFADDVLLSVL 386

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1581	77.4	412	1	MTLR_HUMAN
2	643	31.5	366	1	GHSR_HUMAN
3	641	31.4	364	1	GHSR_RAT
4	641	31.4	366	1	GHSR_PIG
5	614	30.1	257	1	GHSR_MOUSE
6	343	16.8	424	1	NTR1_RAT
7	340	16.6	424	1	NTR1_MOUSE
8	330	16.2	418	1	NTR1_HUMAN
9	274.5	13.4	410	1	NTR2_HUMAN
10	268.5	13.1	416	1	NTR2_RAT
11	268	13.1	454	1	GASR_BOVIN
12	267.5	13.1	370	1	OPRX_CAVPO
13	266	13.0	417	1	NTR2_MOUSE
14	262	12.8	362	1	SSR5_MOUSE
15	254.5	12.5	370	1	OPRX_PIG
16	253.5	12.4	367	1	OPRX_MOUSE
17	253.5	12.4	367	1	OPRX_RAT
18	253.5	12.4	370	1	OPRX_HUMAN
19	253	12.4	453	1	CCNR_XENLA
20	251.5	12.3	473	1	BIAR_CANFA
21	251	12.3	453	1	GASR_MOUSE
22	250.5	12.3	363	1	SSR5_RAT
23	249.5	12.2	452	1	GASR_RABIT
24	249.5	12.2	452	1	GASR_RAT
25	248.5	12.2	360	1	GF25_HUMAN
26	248.5	12.2	418	1	SSR3_HUMAN
27	247.5	12.1	452	1	A2AA_BOVIN
28	247.5	12.1	453	1	GP39_HUMAN
29	247	12.1	453	1	GASR_CANFA
30	245.5	12.0	364	1	SSR5_HUMAN
31	243	11.9	368	1	GALT_HUMAN
32	241.5	11.8	477	1	BIAR_HUMAN
33	241	11.8	370	1	GALT_RAT

34	240	11.7	450	1	GASR_PRANA
35	239	11.7	372	1	GALS_RAT
36	239	11.7	405	1	B3AR_CANFA
37	237.5	11.6	371	1	GALS_MOUSE
38	237.5	11.6	447	1	GASR_HUMAN
39	237.5	11.6	519	1	TLR2_DROME
40	236	11.6	289	1	SSRL_FUGRU
41	233.5	11.4	480	1	BIAR_MACMU
42	233	11.4	395	1	TRPR_CHICK
43	232.5	11.4	351	1	B3AR_CAVPO
44	232.5	11.4	379	1	GRE2_BALAM
45	232.5	11.4	450	1	A2AA_MOUSE

ALIGNMENTS

RESULT 1					
MTLR_HUMAN	MTLR_HUMAN	STANDARD;	PRT;	412 AA.	
AC	O43193;				
DT	15-DEC-1998 (Rel. 37, Created)				
DT	15-DEC-1998 (Rel. 37, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Motilin receptor (G protein-coupled receptor GPR38).				
GN	GPR38 OR MTLR1 OR MTLR.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A. (ISOFORM A).				
RX	MEDLINE=98110578; PubMed=9441746;				
RA	McKee K.K., Tan C.P., Palyha O.C., Liu J., Feighner S.D.,				
RA	Hreniuk D.L., Smith R.G., Howard A.D., van der Ploeg L.H.T.;				
RT	"Cloning and characterization of two human G protein-coupled receptor genes (GPR38 and GPR39) related to the growth hormone secretagogue and neurotensin receptors."				
RT	Genomics 46:426-434(1997).				
RN	[2]				
RP	SEQUENCE FROM N.A. (ISOFORMS A AND B).				
RX	MEDLINE=99316084; PubMed=10381885;				
RA	Feighner S.D., Tan C.P., McKee K.K., Palyha O.C., Hreniuk D.L.,				
RA	Pong S.-S., Austin C.P., Figueroa D., MacNeil D., Cascieri M.A.,				
RA	Nargund R., Bakshi R., Abramovitz M., Stocco R., Katgman S.,				
RA	O'Neill G., van Der Ploeg L.H.T., Evans J., Patchett A.A., Smith R.G.,				
RA	Howard A.D.;				
RT	"Receptor for motilin identified in the human gastrointestinal system."				
RT	Science 284:2184-2188(1999).				
RN	[3]				
RP	SEQUENCE FROM N.A. (ISOFORM A).				
RA	Wall M.;				
RL	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.				
RN	[4]				
RP	FUNCTION.				
RX	MEDLINE=21219832; PubMed=11322507;				
RA	Smith R.G., Leonard R., Bailey A.R.T., Palyha O.C., Feighner S.D.,				
RA	Tan C.P., McKee K.K., Pong S.-S., Griffin P.R., Howard A.D.;				
RT	"Growth hormone secretagogue receptor family members and ligands."				
RL	Endocrine 14:9-14(2001).				
CC	-1- FUNCTION: Receptor for motilin.				
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.				
CC	-1- ALTERNATIVE PRODUCTS:				
CC	Event=Alternative splicing; Named isoforms=2;				
CC	Name=A;				
CC	Isoid=O43193-1; Sequence=Displayed;				
CC	Name=B;				
CC	Isoid=O43193-2; Sequence=VSP_001894;				
CC	-1- TISSUE SPECIFICITY: EXPRESSED ONLY IN THYROID, STOMACH, AND BONE				
CC	NARROW.				
CC	-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.				
CC	-----				

P30796	praoxys nat
O08726	rattus norv
O02662	canis famil
O88854	mus musculu
P32239	homo sapien
P30975	drosophila
O42179	fugu rubrip
P47899	macaca mula
O93603	gallus gall
O60483	cavia porce
O93127	balanus amp
Q01338	mus musculu


```

DR EMBL; U60179; AAC50653.1; -.
DR EMBL; U60181; AAC50654.1; -.
DR EMBL; AF369786; AAK71539.1; -.
DR EMBL; AF369786; AAK71540.1; -.
DR Genbank; HGNC:4267; GHSR.
DR MIM; 601898; -.
DR GO; GO:0016021; C: integral to membrane; TAS.
DR GO; GO:0005886; C: plasma membrane; TAS.
DR GO; GO:0007186; P: G-protein coupled receptor protein signalin. . .; TAS.
DR InterPro; IPR000276; GPCR_Rhodopsin.
DR PRINTS; PR00237; GPCRHOOPS.
DR PROSITE; PS00237; G-PROTEIN RECP F1_1; 1.
DR PROSITE; PS0262; G-PROTEIN RECP F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Alternative splicing.
FT DOMAIN 1 40 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 41 66 1 (POTENTIAL).
FT DOMAIN 67 72 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 73 96 2 (POTENTIAL).
FT DOMAIN 97 117 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 118 139 3 (POTENTIAL).
FT DOMAIN 140 162 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 163 183 4 (POTENTIAL).
FT DOMAIN 184 211 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 212 235 5 (POTENTIAL).
FT DOMAIN 236 263 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 264 285 6 (POTENTIAL).
FT DOMAIN 286 302 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 303 326 7 (POTENTIAL).
FT DOMAIN 327 366 CYTOPLASMIC (POTENTIAL).
FT DISULFID 116 198 BY SIMILARITY.
FT CARBOHYD 13 13 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 266 289 AVVFAFALCWLPFHVGRVLFSSK -> GGSQALRLSLAG
PIISLCLPLSL (in isoform 1B).
/FTID=VSP_001916.
FT VARSPLIC 290 366 Missing (in isoform 1B).
/FTID=VSP_001917.
SQ SEQUENCE 366 AA; 41328 MW; D1B62710DA9DC0C6 CRC64;
Query Match 31.5%; Score 643; DB 1; Length 366;
Best Local Similarity 45.1%; Pred. No. 2.6e-37;
Matches 138; Conservative 43; Mismatches 73; Indels 52; Gaps 7;
QY 5 WNGSDGEA-----REPPWALPPCD---ERRCSPPGLGALVPTAVCLFVWGSGN 56
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 2 WNAITPSEPGFNLTADLDWDASFGNDSGLGDLQLFPAPLLAGVTATCVALFVWGAGN 61
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 57 VVTVMILGRYRDMRTTNTNLVYLSGMAVSDLLILGLPDLVRLWRSRPWFVGPILLCRLSLY 116
DB 62 LLTNLVVSRRPRERTTNTNLVSSNAFSDLLIFLCPLDVLRLWQYRPWNFGDLLCKLFQF 121
QY 117 VBGCTYATLLHMTALSVRYLAICRPLRVLVTRRRVRALIAVLWAVALLSAGPFLFL 176
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 122 VSECTVATLTITALSVERVFAICFPFLAKVVTKGRVKLVIFVIAVAFCSAGPIFVL 181
QY 177 VGVEDQPCISVVGCLNTARIASSPLASSPLWLSRAPPPSPGPTAAALFSRECR 236
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 182 VGVHE-----NGT-----DP--W-----DTNECR 199
QY 237 PS--PAQLGALRVLMTVTTAVFFLPFLCLSLYGLIGRELWSSRRPLRGPAAQRGRGRH 294
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 200 PTFEAVRGLTVMVWSSIFFFLVFLVCLTVLSLIGKLRWRRRGDVAVGASLRDQNHK 259
QY 295 QTVRVL 300
DB : : : : :
DB 260 QTVKML 265
RESULT 3
GHSR_RAT
ID _GHSR_RAT STANDARD; PRT; 364 AA.

```

```

AC 008725;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Growth hormone secretagogue receptor type 1 (GHS-R) (GH-releasing
DE peptide receptor) (GHRP) (Ghrelin receptor).
GN GHSR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Pituitary;
RX MEDLINE=9724655; PubMed=9092793;
RA McKee K.K., Palyha O.C., Feighner S.D., Hreniuk D.L., Tan C.P.,
RA Phillips M.S., Smith R.G., der Ploeg L.H.T., Howard A.D.;
RT "Molecular analysis of rat pituitary and hypothalamic growth hormone
RT secretagogue receptors.";
RL Mol. Endocrinol. 11:415-423(1997).
RN [2]
RP SEQUENCE OF 1-240 FROM N.A.
RC STRAIN=Wistar; TISSUE=Pituitary;
RX MEDLINE=98100386; PubMed=9437732;
RA Yokote R., Sato M., Matsubara S., Ohye H., Niimi M., Murao K.,
RA Takahara J.;
RT "Molecular cloning and gene expression of growth hormone-releasing
RT peptide receptor in rat tissues.";
RL Peptides 19:15-20(1998).
RN [3]
RP FUNCTION.
RX MEDLINE=20067959; PubMed=10604470;
RA Kojima M., Hosoda H., Date Y., Nakazato M., Matsuo H., Kangawa K.;
RT "Ghrelin is a growth-hormone-releasing acylated peptide from
RT stomach.";
RL Nature 402:656-660(1999).
CC -1- FUNCTION: Receptor for ghrelin, coupled to G-alpha-11 proteins.
CC Stimulates growth hormone secretion. Binds also other growth
CC hormone releasing peptides (GHRP) (e.g. Met-enkephalin and GHRP-6)
CC as well as non-peptide, low molecular weight secretagogues (e.g.
CC L-692,429, MK-0677, adenosine) (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
-----
DR EMBL; U94321; AAC33156.1; -.
DR EMBL; AB001982; BAA21777.1; ALT INT.
DR InterPro; IPR000276; GPCR_Rhodopsin.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCRHOOPS.
DR PROSITE; PS00237; G-PROTEIN RECP F1_1; 1.
DR PROSITE; PS0262; G-PROTEIN RECP F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 40 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 41 66 1 (POTENTIAL).
FT DOMAIN 67 72 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 73 96 2 (POTENTIAL).
FT DOMAIN 97 117 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 118 139 3 (POTENTIAL).
FT DOMAIN 140 162 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 163 183 4 (POTENTIAL).
FT DOMAIN 184 211 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 212 235 5 (POTENTIAL).
FT DOMAIN 236 263 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 264 285 6 (POTENTIAL).
FT DOMAIN 286 302 EXTRACELLULAR (POTENTIAL).

```

```
FT TRANSMEM 303 326 7 (POTENTIAL).
FT DOMAIN 327 364 CYTOPLASMIC (POTENTIAL).
FT DISULFID 115 197 BY SIMILARITY.
FT CARBOHYD 13 13 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 26 26 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 364 AA; 40963 MW; DCBF59B061EEB9 CRC64;

Query Match 31.4%; Score 641; DB 1; Length 364;
Best Local Similarity 45.8%; Pred. No. 3.6e-37;
Matches 142; Conservative 41; Mismatches 65; Indels 62; Gaps 10;

Qy 5 WNGSDGPEGAREP-----PWPALPPCD-----ERRCSPFFPLGALVPVAVCLCLFVVGSG 55
Db 2 WNAT--EPEEPNVLTDLDWDASGNDLSPDELLPLFPAPLLAGVTATCVALFVVGISG 59
Qy 56 NVVTMLIGRYDRMTTNNLYLSGMVSDLLILGLFDPDLRLWRSPWPFGPLCLRLSL 115
Db 60 NLLTMLVVSFRPREMTTNNLYLSGMVSDLLILGLFDPDLRLWRSPWPFGPLCLRLSL 119
Qy 116 YVGGCTYATLLHMTALSVRYLAICRPLARVLVTRRRVRLAVLWAVALLSAGPFLF 175
Db 120 FVSECTYATVLTITALLSVRYLAICRPLARVLVTRRRVRLAVLWAVALLSAGPFLF 179
Qy 176 LVGVEQDPSVWVGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAEAAALFSREC 235
Db 180 LVGVEHE-----NGT-----DPRD-----TNEC 197
Qy 236 RPS--PAQLGALRVMLWVTTAYFELPFLCLSLILYGLIGRELWSSRRPLRGPAASG---RE 290
Db 198 RATEFAVRSGLLTVMWVSSVFFFLPVFCLTVLSLIGRKLW---RR--RGDAVAGSLRD 253
Qy 291 RGHRTQVRVL 300
Db 254 QNHQTVKML 263

RESULT 4
ID_GHSR_PIG STANDARD; PRT; 366 AA.
AC Q95254; Q95255;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Growth hormone secretagogue receptor type 1 (GHS-R) (GH-releasing
DE peptide receptor) (GHRP) (Ghrelin receptor).
GN GHSR.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1A AND 1B).
RC STRAIN=Yorkshire; TISSUE=Pituitary;
RX MEDLINE=96337998; PubMed=8688086;
RA Howard A.D., Feighner S.D., Cully D.F., Hamelin M., Hreniuk D.L., Liu K.K.,
RA Liberator P.A., Rosenblum C.I., Pares P.S., Diaz C., Chou M., Liu K.K.,
RA Palyha O.C., Anderson J., Pares P.S., Diaz C., Chou M., Liu K.K.,
RA McKee K.K., Pong S.-S., Chang L.-Y., Elbrecht A., Dashkevich M.,
RA Heavens R., Rigby M., Sirinathsinghji D.J.S., Dean D.C., Melillo D.G.,
RA Patchett A.A., Nargund R., Griffin P.R., Demartino J.A., Gupta S.K.,
RA Schaeffer J.M., Smith R.G., van der Ploeg L.H.T.;
RT "A receptor in pituitary and hypothalamus that functions in growth
RT hormone release."
RL Science 273:974-977(1996).
CC -1- FUNCTION: Receptor for ghrelin, coupled to G-alpha-11 proteins.
CC Stimulates growth hormone secretion. Binds also other growth
CC hormone releasing peptides (GHRP) (e.g. Met-enkephalin and GHRP-6)
CC as well as non-peptide, low molecular weight secretagogues (e.g.
CC L-692,429, MK-0677, adenosine).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1A;
```

```
CC CC IsoId=Q95254-1; Sequence=Displayed;
CC CC Name=1B;
CC CC IsoId=Q95254-2; Sequence=VSP 001918, VSP 001919;
CC CC -1- TISSUE SPECIFICITY: PITUITARY AND HYPOTHALAMUS.
CC CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC CC
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC
CC CC EMBL; U60178; AAC48630.1; -
CC CC EMBL; U60180; AAC48631.1; -
CC CC InterPro; IPR000276; GPCR_Rhodopsin.
CC CC Pfam; PF00001; 7tm.1; 1.
CC CC PRINTS; PR00237; GPCRRHODOPSIN.
CC CC PROSITE; PS00237; G PROTEIN RECP FL 1; 1.
CC CC PROSITE; PS00262; G-PROTEIN RECP FL 2; 1.
CC CC G-protein coupled receptor; Transmembrane; Glycoprotein;
CC CC Alternative splicing.
CC CC DOMAIN 1 40 EXTRACELLULAR (POTENTIAL).
CC CC TRANSMEM 41 66 1 (POTENTIAL).
CC CC DOMAIN 67 72 CYTOPLASMIC (POTENTIAL).
CC CC TRANSMEM 73 96 2 (POTENTIAL).
CC CC DOMAIN 97 117 EXTRACELLULAR (POTENTIAL).
CC CC TRANSMEM 118 139 3 (POTENTIAL).
CC CC DOMAIN 140 162 CYTOPLASMIC (POTENTIAL).
CC CC TRANSMEM 163 183 4 (POTENTIAL).
CC CC DOMAIN 184 211 EXTRACELLULAR (POTENTIAL).
CC CC TRANSMEM 212 235 5 (POTENTIAL).
CC CC TRANSMEM 236 263 6 (POTENTIAL).
CC CC TRANSMEM 264 285 7 (POTENTIAL).
CC CC TRANSMEM 303 326 7 (POTENTIAL).
CC CC DOMAIN 327 366 BY SIMILARITY.
CC CC DISULFID 116 198 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC CARBOHYD 13 13 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC CARBOHYD 27 27 AVVFAFILCLWPHVGRYLFSSK -> GGSQCALELSLP
CC CC VARSPLIC 266 PLHSSCLFSSP (in isoform 1B).
CC CC VARSPLIC 290 Missing (in isoform 1B).
CC CC VARSPLIC 366 /FTID=VSP 001919.
CC CC SEQUENCE 366 AA; 41194 MW; 2C850B3EF61B7C1C CRC64;
CC CC
CC CC Query Match 31.4%; Score 641; DB 1; Length 366;
CC CC Best Local Similarity 45.3%; Pred. No. 3.6e-37;
CC CC Matches 140; Conservative 41; Mismatches 70; Indels 58; Gaps 8;
Qy 5 WNGSDGPEGA-----REPPWALPPCD-----ERRCSPFFPLGALVPVAVCLCLFVVGSGN 56
Db 2 WNATSEEPGNLTLPDLGWDAPPENDSLVEELLPLFPTELLAGVTATCVALFVVGIAGN 61
Qy 57 VVTVMILIGRYDRMTTNNLYLSGMVSDLLILGLFDPDLRLWRSPWPFGPLCLRLSL 116
Db 62 LLLTMLVVSFRPREMTTNNLYLSGMVSDLLILGLFDPDLRLWRSPWPFGPLCLRLSL 121
Qy 117 VGGCTYATLLHMTALSVRYLAICRPLARVLVTRRRVRLAVLWAVALLSAGPFLF 176
Db 122 VSECTYATVLTITALLSVRYLAICRPLARVLVTRRRVRLAVLWAVALLSAGPFLF 181
Qy 177 VGVEQDPSVWVGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAEAAALFSREC 236
Db 182 VGVEHD-----NGT-----DPRD-----TNEC 199
Qy 237 PS--PAQLGALRVMLWVTTAYFELPFLCLSLILYGLIGRELWSSRRPLRGPAASG---RER 291
Db 200 ATEFAVRSGLLTVMWVSSVFFFLPVFCLTVLSLIGRKLW---RRRGRGAAGVSSLRDQ 256
Qy 292 GHRQTVRVL 300
```



```

Db 257 NHKQTVKWL 265
|:|:|:|:|
Best Local Similarity 45.1%; Pred. No. 1.8e-35;
Matches 137; Conservative 39; Mismatches 66; Indels 62; Gaps 10;

QY 5 WNGSDGEGAREP-----PWPALPPCD---ERRCSPPFPGALVPVTVAVCLCLFVVVGVS 55
DB 2 WNAAT--PSEPEPNTVLDLDWDASFGNDLSDELLPLFPAPLLAGVATATCAVALFVVGIS 59
QY 56 NVVTVMILGRYDMRTTNLYLGSMVAVSOLLILLGLPFDLYRLWRSRPVVFGPLLCRLSL 115
DB 60 NLLTLMVWSRFRRLRTTNLYLSSMAFSDLLIFLCPDLVRLQVYRPNWFGDLCKLQ 119
QY 116 YVCEGCTATLLHMTALSVERVLAICPLARARVLTTRRRVRLIALVLMVALLSAGPPLF 175
DB 120 FVSESCYATVLTITALSVERVFAICPLRAKVVVTKGRVKLVILVAVAFCSAGPIFV 179
QY 176 LVGVEQDPGISVVPGLNGTARTIASSPLSPPLWLSRAPPPSPSGPETAEEAALFSREC 235
DB 180 LVGVEHE-----NGT-----DPRD-----TNEC 197
QY 236 RPS--PAQLGALRVMLWVTTAYFPFLCLSLYLGILGRELMSRRPLRGAASG---RE 290
DB 198 RATEPAVRSGLLTVMVWVSVFPFLPVFCLTVLYSLIGRKLW--RR--RGDAAVGSSLRD 253
QY 291 RGRH 294
DB 254 QNHK 257

RESULT 6
NTRL_RAT
ID NTRL_RAT STANDARD; PRT; 424 AA.
AC P20789;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Neurotensin receptor type 1 (NT-R-1) (High-affinity leucobastine-
DE insensitive neurotensin receptor) (NTRH).
GN NTRH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=90297956; PubMed=1694443;
RA Tanaka K., Masu M., Nakanishi S.;
RT "Structure and functional expression of the cloned rat neurotensin
RT receptor.";
RL Neuron 4:847-854 (1990).
CC -1- FUNCTION: RECEPTOR FOR THE TRIPEPTIDE NEUROTENSIN. IT IS
CC ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
CC CALCIUM SECOND MESSENGER SYSTEM.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC HIGHEST TO TACHYKININS RECEPTORS.
DR PIR; JH0164; JH0164.
DR InterPro; IPR000276; GPCR_Rhodopsin.
DR InterPro; IPR003985; NTL_rec.
DR InterPro; IPR003984; NTL_rec.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCRRHODOPSIN.
DR PRINTS; PR01479; NEUROTENSINR.
DR PRINTS; PR01480; NEUROTENSINR.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL_1; 1.
DR PROSITE; PS0262; G-PROTEIN RECEPTOR FL_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW phosphorylation; Lipoprotein; Palmitate.
FT DOMAIN 1 64 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 65 87 1 (POTENTIAL).
FT DOMAIN 88 96 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 97 121 2 (POTENTIAL).
FT DOMAIN 122 143 EXTRACELLULAR (POTENTIAL).

Db 257 NHKQTVKWL 265
|:|:|:|:|
Best Local Similarity 45.1%; Pred. No. 1.8e-35;
Matches 137; Conservative 39; Mismatches 66; Indels 62; Gaps 10;

QY 5 WNGSDGEGAREP-----PWPALPPCD---ERRCSPPFPGALVPVTVAVCLCLFVVVGVS 55
DB 2 WNAAT--PSEPEPNTVLDLDWDASFGNDLSDELLPLFPAPLLAGVATATCAVALFVVGIS 59
QY 56 NVVTVMILGRYDMRTTNLYLGSMVAVSOLLILLGLPFDLYRLWRSRPVVFGPLLCRLSL 115
DB 60 NLLTLMVWSRFRRLRTTNLYLSSMAFSDLLIFLCPDLVRLQVYRPNWFGDLCKLQ 119
QY 116 YVCEGCTATLLHMTALSVERVLAICPLARARVLTTRRRVRLIALVLMVALLSAGPPLF 175
DB 120 FVSESCYATVLTITALSVERVFAICPLRAKVVVTKGRVKLVILVAVAFCSAGPIFV 179
QY 176 LVGVEQDPGISVVPGLNGTARTIASSPLSPPLWLSRAPPPSPSGPETAEEAALFSREC 235
DB 180 LVGVEHE-----NGT-----DPRD-----TNEC 197
QY 236 RPS--PAQLGALRVMLWVTTAYFPFLCLSLYLGILGRELMSRRPLRGAASG---RE 290
DB 198 RATEPAVRSGLLTVMVWVSVFPFLPVFCLTVLYSLIGRKLW--RR--RGDAAVGSSLRD 253
QY 291 RGRH 294
DB 254 QNHK 257

RESULT 6
NTRL_RAT
ID NTRL_RAT STANDARD; PRT; 424 AA.
AC P20789;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Neurotensin receptor type 1 (NT-R-1) (High-affinity leucobastine-
DE insensitive neurotensin receptor) (NTRH).
GN NTRH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=90297956; PubMed=1694443;
RA Tanaka K., Masu M., Nakanishi S.;
RT "Structure and functional expression of the cloned rat neurotensin
RT receptor.";
RL Neuron 4:847-854 (1990).
CC -1- FUNCTION: RECEPTOR FOR THE TRIPEPTIDE NEUROTENSIN. IT IS
CC ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
CC CALCIUM SECOND MESSENGER SYSTEM.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC HIGHEST TO TACHYKININS RECEPTORS.
DR PIR; JH0164; JH0164.
DR InterPro; IPR000276; GPCR_Rhodopsin.
DR InterPro; IPR003985; NTL_rec.
DR InterPro; IPR003984; NTL_rec.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCRRHODOPSIN.
DR PRINTS; PR01479; NEUROTENSINR.
DR PRINTS; PR01480; NEUROTENSINR.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL_1; 1.
DR PROSITE; PS0262; G-PROTEIN RECEPTOR FL_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW phosphorylation; Lipoprotein; Palmitate.
FT DOMAIN 1 64 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 65 87 1 (POTENTIAL).
FT DOMAIN 88 96 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 97 121 2 (POTENTIAL).
FT DOMAIN 122 143 EXTRACELLULAR (POTENTIAL).

Query Match
30.1%; Score 614; DB 1; Length 257;

```

FT TRANSMEM 144 165 3 (POTENTIAL).
 FT DOMAIN 166 188 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 189 210 4 (POTENTIAL).
 FT DOMAIN 211 235 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 236 260 5 (POTENTIAL).
 FT DOMAIN 261 308 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 309 330 6 (POTENTIAL).
 FT DOMAIN 331 348 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 349 372 7 (POTENTIAL).
 FT DOMAIN 373 424 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 4 4 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 38 38 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 42 42 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT DISULFID 142 225 BY SIMILARITY.
 FT LIPID 388 PALMITATE (POTENTIAL).
 SQ SEQUENCE 424 AA; 47054 MW; A9C2P7EAF89BCD3 CRC64;

Query Match 16.8%; Score 343; DB 1; Length 424;
 Best Local Similarity 33.9%; Pred. No. 1.1e-16;
 Matches 82; Conservative 41; Mismatches 71; Indels 48; Gaps 6;

QY 39 VPVTAVCLCLFVVGSGNVVTMLIGR---YRDMRTTNLYLGSMAVSDLLI-LLGLPFD 94
 DB 65 VLVTAYLALFVVGTVGNSVTAFTLARKKSLSQSTVHYHLSGLSALLILLAMPVE 124
 QY 95 LYSR-LWRSRPVWFGPILLCRSLYVGGCTYATLLHMTALSVERYLAICRPLRVLVTRR 153
 DB 125 LYNFIWVHHPWAFGDACRGYFLRDCTATYATNVALSVSLVERYLAICHFPKATLMRS 184
 QY 154 RVRLATVAVLWALLSAGPFLFLVGVQDPGIVSVVGLNGTARIASSPLASSPPLMSRA 213
 DB 185 RTKKFISAIWLASALLAIPMLFTWGLQNRSGDGTGHPG-----GLVCTPIVDT----- 231
 QY 214 PPPSPSGPTAEAAALFSRECRPSPAQLGALRVMLVWTAYFEL-PFLCLSLYLGLIGR 272
 DB 232 -----ATVKVVIQVNTFMSFLPMLVISILNTVIAN 262
 QY 273 EL 274
 DB 263 KL 264

RESULT 7
 NTRI MOUSE STANDARD; PRT; 424 AA.
 AC O88319; 2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neurotensin receptor type 1 (NT-R-1).
 GN NTSR1 OR NTSR.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Snider J., Sano H., Ohta M.;
 RT Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: RECEPTOR FOR THE TRIPEPTIDE NEUROTENSIN. IT IS
 CC ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
 CC CALCIUM SECOND MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC HIGHEST TO TACHYKININS RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB017027; BAA33013.1; --
 CC MGI; 97386; Nscr.
 CC InterPro; IPR000276; GPCR Rhodopsin.
 CC InterPro; IPR003985; NTL rec.
 CC InterPro; IPR003984; NTL rec.
 CC Pfam; PF00001; 7tm.1.1_
 CC PRINTS; PR00237; GPCRHOPOPSN.
 CC PRINTS; PR01479; NEUROTENSINR.
 CC PRINTS; PR01480; NEUROTENSINR.
 CC PROSITE; PS00237; G PROTEIN RECEPTOR F1.1; 1.
 CC PROSITE; PS0262; G PROTEIN RECEPTOR F1.2; 1.
 CC G-protein coupled receptor; Transmembrane; Glycoprotein;
 CC phosphorylation; Lipoprotein; Palmitate.
 FT DOMAIN 1 64 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 65 87 1 (POTENTIAL).
 FT DOMAIN 88 96 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 97 121 2 (POTENTIAL).
 FT DOMAIN 122 143 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 144 165 3 (POTENTIAL).
 FT DOMAIN 166 188 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 189 210 4 (POTENTIAL).
 FT DOMAIN 211 235 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 236 260 5 (POTENTIAL).
 FT DOMAIN 261 308 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 309 330 6 (POTENTIAL).
 FT DOMAIN 331 348 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 349 372 7 (POTENTIAL).
 FT DOMAIN 373 424 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 4 4 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 38 38 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 42 42 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT DISULFID 141 224 BY SIMILARITY.
 FT LIPID 388 PALMITATE (POTENTIAL).
 SQ SEQUENCE 424 AA; 47216 MW; 8E9A723171A48711 CRC64;

Query Match 16.8%; Score 340; DB 1; Length 424;
 Best Local Similarity 34.7%; Pred. No. 1.7e-16;
 Matches 84; Conservative 39; Mismatches 71; Indels 48; Gaps 7;

QY 39 VPVTAVCLCLFVVGSGNVVTMLIGR---YRDMRTTNLYLGSMAVSDLLI-LLGLPFD 94
 DB 64 VLVTAYLALFVVGTVGNSVTAFTLARKKSLSQSTVHYHLSGLSALLILLAMPVE 123
 QY 95 LYSR-LWRSRPVWFGPILLCRSLYVGGCTYATLLHMTALSVERYLAICRPLRVLVTRR 153
 DB 124 LYNFIWVHHPWAFGDACRGYFLRDCTATYATNVALSVSLVERYLAICHFPKATLMRS 183
 QY 154 RVRLATVAVLWALLSAGPFLFLVGVQDPGIVSVVGLNGTARIASSPLASSPPLMSRA 213
 DB 184 RTKKFISAIWLASALLAIPMLFTWGLQ-----
 QY 214 PPPSPSGPTAEAAALFSRECRPSPAQLGALRVMLVWTAYFEL-PFLCLSLYLGLIGR 272
 DB 214 ADAQHPGG-----LVCTPT-VDTATVKVVIQVNTFMSFLPMLVISILNTVIAN 261
 QY 273 EL 274
 DB 262 KL 263

RESULT 8
 NTRI HUMAN STANDARD; PRT; 418 AA.
 AC P30959; Q9H4H1; Q9H4T5;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Neurotensin receptor type 1 (NT-R-1) (High-affinity leucocastine-
 DE insensitive neurotensin receptor) (NTRH).

NTSR1 OR NTRR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=93154505; PubMed=8381365;
 RA Vita N., Laurent P., Lefort S., Chalou P., Dumont X., Kaghad M.,
 RA Gully D., le Fur G., Ferrara P., Caput D.;
 RT "Cloning and expression of a complementary DNA encoding a high
 RT affinity human neurotensin receptor.";
 RL FEBS Lett. 317:139-142(1993).
 RN [2]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=21638749; PubMed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.V., Clee C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Leharasaiho M.H., Leverhale M.A., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McConnachie L.J., McKay K., McMurray A.A.,
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulten J.E.,
 RA Swann R.M., Symamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20.";
 RL Nature 414:865-871(2001).
 CC -1- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUROTENSIN. IT IS
 CC ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
 CC CALCIUM SECOND MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC HIGHEST TO TACHYKINS RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X70070; CAA49675.1; -;
 DR EMBL; ALJ57033; CAC14923.1; -;
 DR EMBL; AL035669; CAC12747.1; -;
 DR PIR; S29506; S29506.
 DR HSP; P02699; IP88.
 DR Gnew; HNCN:8039; NTSR1.
 DR MIM; 162651; -;
 DR GO; GO:0005783; C:endoplasmic reticulum; TAS.
 DR GO; GO:0005794; C:Golgi apparatus; TAS.
 DR GO; GO:0005987; C:integral to plasma membrane; TAS.
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; TAS.
 DR GO; GO:0007268; P:synaptic transmission; TAS.
 DR InterPro; IPR00276; GPCR_Rhodopsin.
 DR InterPro; IPR003985; NTL1 Rec.
 DR InterPro; IPR003984; NTL1 Rec.

DR PFam; PF00001; 7tm1.1; 1.
 DR PRINTS; PRO0237; GPCRHHODPSN.
 DR PRINTS; PRO1479; NEUROTENSINR.
 DR PRINTS; PRO1480; NEUROTENSINR.
 DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; 1.
 DR PROSITE; PS0262; G_PROTEIN_RECP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW Phosphorylation; Lipoprotein; Palmitate.
 FT DOMAIN 1 63 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 64 86 1 (POTENTIAL).
 FT DOMAIN 87 95 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 96 120 2 (POTENTIAL).
 FT DOMAIN 121 142 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 143 164 3 (POTENTIAL).
 FT DOMAIN 165 187 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 188 209 4 (POTENTIAL).
 FT DOMAIN 210 234 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 235 259 5 (POTENTIAL).
 FT DOMAIN 260 303 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 304 325 6 (POTENTIAL).
 FT DOMAIN 326 343 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 344 367 7 (POTENTIAL).
 FT DOMAIN 368 418 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 4 4 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 37 37 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 41 41 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT DISULFID 141 224 BY SIMILARITY.
 FT LIPID 383 393 PALMITATE (POTENTIAL).
 FT CONFLICT 200 200 T -> A (IN REF. 2).
 SQ SEQUENCE 418 AA; 46288 MW; BBD1EBC2BE6E390 CRC64;
 Query Match 16.2%; Score 330; DB 1; Length 418;
 Best Local Similarity 31.2%; Pred. No. 8.3e-16;
 Matches 89; Conservative 44; Mismatches 88; Indels 64; Gaps 9;
 QY 39 VPVTAVCLCFVVGSGNVVTVNLIGR---YRDMRTTNYLGSMVSDLL-ILLGLPPD 94
 DB 64 VLVTVYALFVVGTVGNTVTAFTLARKKLSQSLQSTVHYHGLSLALSLLTLLAMPVE 123
 QY 95 LYR-LWRSPWVPGPLLCRLSLYVGGCTVATLLHMTALSVRYLAICRPLRARVLVTR 153
 DB 124 LYNFIVHHHPAFGDAGCGYFLRDACVATYALNVASUSVERLYLAICHPFKAKTLMSRS 183
 QY 154 RYRALJAVLWAVALLSAGPFLVLGVQDPGIVSVPLNGTARIASSPLLSRA 213
 DB 184 RYKFFISALWLASALITVPLFTMG-EQN----- 211
 QY 214 PPSPSPGPETAALFSPRECRPSPAQLGALRVMLWVTT-AVFPFLCLSLYLIGLR 272
 DB 212 -----RSADGQHAGGLVCTPT-IHTATVKVVIQVNTFMSFIPPMVVISLVNTIAN 261
 QY 273 ELWSSRRPL--RQPAAS-----GREGRHQTQTVRLR 301
 DB 262 KLTVMYVQAQBOGQVCTVGEHSTFSMAIEPGRVQALRHGVRLR 306
 RESULT 9
 NTR2 HUMAN STANDARD; PRT; 410 AA.
 AC O95665; Q8TBH6;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Neurotensin receptor type 2 (NT-R-2) (Levocastine-sensitive
 DE neurotensin receptor) (NTR2 receptor).
 GN NTSR2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Brain;

RX MEDLINE=99066919; PubMed=9851594;
 RA Vita N., Oury-Donat F., Chalton P., Guillemot M., Kaghad M., Bachy A.,
 RA Thurneysen S., Garcia S., Poinet-Chazel C., Casellas P., Keane P.,
 RA Le Fur G., Maffrand J.-P., Shoubrie P., Caput D., Ferrara P.,
 RT "Neurotensin is an antagonist of the human neurotensin NT2 receptor
 RL expressed in Chinese hamster ovary cells";
 RN Eur. J. Pharmacol. 360:265-272(1998).
 RP [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haie F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Tohiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Heiton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitting R.W., Touchman J.W., Green E.D., Bouffard G.G.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUROTENSIN. IT IS
 CC ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
 CC CALCIUM SECOND MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC HIGHEST TO TACHYKININS RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Y10148; CAA71233.1; --
 DR EMBL; BC022501; AAH22501.1; --
 DR Genew; HGNC:8040; NTSR2.
 DR MIM; 605538; --
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.
 DR GO; GO:0007600; P:sensory perception; TAS.
 DR InterPro; IPR000276; GPCR Rhodops.
 DR InterPro; IPR003986; NT2 rec.
 DR InterPro; IPR003984; NT rec.
 DR Pfam; PF00001; 7tm_1; 1
 DR PRINTS; PR00237; GPCRHOOPS.
 DR PRINTS; PR01479; NEUROTENSIN.
 DR PRINTS; PR01481; NEUROTENSIN2.
 DR PROSITE; PS00237; G PROTEIN RECP FL 1; 1.
 DR PROSITE; PS0262; G PROTEIN RECP FL 2; 1.
 KW G-protein coupled receptor; Transmembrane; Lipoprotein; Palmitate.
 FT DOMAIN 1 32 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 33 55 1 (POTENTIAL).
 FT DOMAIN 56 64 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 65 87 2 (POTENTIAL).
 FT DOMAIN 88 109 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 110 131 3 (POTENTIAL).
 FT DOMAIN 132 154 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 155 176 4 (POTENTIAL).
 FT DOMAIN 177 217 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 218 237 5 (POTENTIAL).
 FT DOMAIN 238 297 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 298 318 6 (POTENTIAL).
 FT DOMAIN 319 337 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 338 358 7 (POTENTIAL).
 FT DOMAIN 359 410 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 108 194 BY SIMILARITY.
 FT LIPID 377 377 PALMITATE (POTENTIAL).
 FT CONFLICT 168 168 L -> M (IN REF. 2).
 FT CONFLICT 367 367 K -> R (IN REF. 2).
 SQ SEQUENCE 410 AA; 45413 MW; 8C3ADA22BE15FD66 CRC64;
 Query Match 13.4%; Score 274.5; DB 1; Length 410;
 Best Local Similarity 32.6%; Pred. No. 5.2e-12;
 Matches 92; Conservative 36; Mismatches 91; Indels 63; Gaps 11;
 QY 17 PWPALPP---CDERCSFPFPGALVPTAVCLCLFVGVSGVNVVTVMLI-----GR 65
 DB 8 PRPSNPGSLDARGVDVTRLWAKVLFALTALVALIWAAGNALSVHVVHVKARAGR 67
 QY 66 YRDMRTTTLVGLSMAVSDLLIL-GLPFDLYR-LWRSRPWFVGPGLLCRLSLVVGEGCTY 123
 DB 68 LRH-----HVLSLALAGLLLVGVVELYSFVWFHYFVFGDLGCRGYFVHLCAY 120
 QY 124 ATLLHMTALSVRYLAICPLRVLVTRRRVRLIAVLMAVALLSAGPPLFLVGV---- 179
 DB 121 ATVLSVAGLSAERCLAVCQPLRARSLLTPRTRVLVLSWAASLGLALPMVIMGKH 180
 QY 180 -----EQDPSISVPGINGTARIAS-----PLASSPPL-----WLS 211
 DB 181 ETADGEPEPASRVCTVL--VSRTALQVFIQNVNLSFVLPLALTAFANGVTVSHLLALCS 238
 QY 212 RAPPSPPPGPGPTAAEAALFSRECRPSAQLGALRVMLVWT 253
 DB 239 QVPSISTP-GSSTPSRLLELSEE-----GLLSFIVWKKY 271
 RESULT 10
 NTR2 RAT
 ID NTR2 RAT STANDARD; PRT; 416 AA.
 AC Q63384;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neurotensin receptor type 2 (NT-R-2) (High-affinity leucobastine-
 DE sensitive neurotensin receptor).
 GN NTSR2 OR NTR2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Hypothalamus;
 RX MEDLINE=96228041; PubMed=8647296;
 RA Chalton P., Vita N., Kaghad M., Guillemot M., Bonin J.,
 RA Delpech B., le Fur G., Ferrara P., Caput D.;
 RT "Molecular cloning of a leucobastine-sensitive neurotensin binding
 RT site";
 RL FEBS Lett. 386:91-94(1996).
 CC -1- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUROTENSIN. IT IS
 CC ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
 CC CALCIUM SECOND MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: ABUNDANT IN CORTEX AND HYPOTHALAMUS, AND LOWER
 CC LEVELS SEEN IN THE HEART AND INTESTINE.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED MAXIMALLY IN 7-DAY-OLD BRAIN AND
 CC EXPRESSION DECREASES PROGRESSIVELY UNTIL ADULTHOOD (35-DAY-OLD
 CC BRAIN).
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC HIGHEST TO TACHYKININS RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X97121; CAA65787.1; -
 FIR; S68822; S68822.
 InterPro: IPR000276; GPCR_Rhodopsin.
 InterPro: IPR003986; NT2_rec.
 InterPro: IPR003984; NT2_rec.
 Pfam: PF00001; 7tm_1; 1.
 PRINTS; PR00237; GPCR_Rhodopsin.
 PRINTS; PR01479; NEUROTENSINR.
 PRINTS; PR01481; NEUROTENSINR.
 PROSITE; PS00237; G PROTEIN RECP F1_1; 1.
 PROSITE; PS00262; G PROTEIN RECP F1_2; 1.
 G-protein coupled receptor; Transmembrane; Lipoprotein; Palmitate.
 FT DOMAIN 1 32
 FT TRANSMEM 33 55
 FT DOMAIN 56 64
 FT TRANSMEM 65 87
 FT DOMAIN 88 109
 FT TRANSMEM 110 131
 FT DOMAIN 132 154
 FT TRANSMEM 155 176
 FT DOMAIN 177 216
 FT TRANSMEM 217 237
 FT DOMAIN 238 297
 FT TRANSMEM 298 318
 FT DOMAIN 319 337
 FT TRANSMEM 338 358
 FT DOMAIN 359 416
 FT DISULFID 108 194
 FT LIPID 377 377
 SEQUENCE 416 AA; 46265 MW; 127C5F5CB8FE208 CRC64;

Query Match 13.1%; Score 268.5; DB 1; Length 416;
 Best Local Similarity 37.9%; Pred. No. 1.4e-11;
 Matches 66; Conservative 33; Mismatches 66; Indels 9; Gaps 4;
 QY 18 PWALP-----PCDERRCSPFPGALVPTAVTAVCLCLFVGVGSGNVVTVMLGRYDRMT 71
 DB 6 PWPFRSPSAGLSLEARGLVDTLWAKVLTALYSILFAFGTAGNALS VHVILKARGRP 65
 QY 72 -TTNLVGLSMAVSDLLTLL-GLPFDLYR-LWRGRPWVFGPLLCRLSLYVGEGCTYATLLH 128
 DB 66 GLRYHVLSLALSALLLLVSMPELYNFVWSHPVFGDLGCRGYFVRELCAVATLS 125
 QY 129 MTALSVERYLAICRPLARVLTTRRRVRLIAVLMAVALLSAGPFLFLVGVEQD 182
 DB 126 VASLSAERCLAVCPLRARELLTPRTRRLSLVWVASLGLALPMAVIMQKHE 179

RESULT 11
 GASK_BOVIN STANDARD; PRT; 454 AA.
 AC P79266;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Gastrin/cholecystokinin type B receptor (CKC-B receptor) (CKC-BR).
 GN CCKBR.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RX MEDLINE=97003869; PubMed=8851180;
 RA Dufréne M., Escrieu C., Clerc P., le Huerou-Luron I., Prats H.,
 Bertrand V., le Meuth V., Guilloteau P., Vaysse N., Fourmy D.,

RT RT Molecular cloning, developmental expression and pharmacological
 RL RT characterization of the CCKR/gastrin receptor in the calf pancreas.;
 CC Eur. J. Pharmacol. 297:165-179(1996).
 CC -1- FUNCTION: RECEPTOR FOR GASTRIN AND CHOLECYSTOKININ. THE CCK-B
 CC RECEPTORS OCCUR THROUGHOUT THE CENTRAL NERVOUS SYSTEM WHERE THEY
 CC MODULATE ANXIETY, ANALGESIA, AROUSAL, AND NEUROLEPTIC ACTIVITY.
 CC THIS RECEPTOR MEDIATES ITS ACTION BY ASSOCIATION WITH G PROTEINS
 CC THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER
 CC SYSTEM.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; S83090; AAB46896.1; -
 HSSP; P02699; 1F88.
 InterPro: IPR000276; GPCR_Rhodopsin.
 Pfam; PF00001; 7tm_1; 1.
 PRINTS; PR00237; GPCR_Rhodopsin.
 PROSITE; PS00237; G PROTEIN RECP F1_1; 1.
 PROSITE; PS00262; G PROTEIN RECP F1_2; 1.
 G-protein coupled receptor; Transmembrane; Glycoprotein;
 Lipoprotein; Palmitate; Phosphorylation.
 FT DOMAIN 1 57
 FT TRANSMEM 58 79
 FT DOMAIN 80 87
 FT TRANSMEM 88 109
 FT DOMAIN 110 131
 FT TRANSMEM 132 150
 FT DOMAIN 151 170
 FT TRANSMEM 171 189
 FT DOMAIN 190 219
 FT TRANSMEM 220 242
 FT DOMAIN 243 340
 FT TRANSMEM 341 362
 FT DOMAIN 363 380
 FT TRANSMEM 381 401
 FT DOMAIN 402 454
 FT CARBOHYD 7 7
 FT CARBOHYD 30 30
 FT CARBOHYD 36 36
 FT DISULFID 127 205
 FT LIPID 415 415
 SEQUENCE 454 AA; 48781 MW; A2846A580508ABA6 CRC64;

Query Match 13.1%; Score 268; DB 1; Length 454;
 Best Local Similarity 25.6%; Pred. No. 1.6e-11;
 Matches 99; Conservative 60; Mismatches 139; Indels 88; Gaps 14;
 QY 2 GSP-WNGSDGPEGAREPPMPALPCDERRCSPFPGALVPTAVTAVCLCLFVGVGSGNVTV 60
 DB 25 GGPLNGSGTGNLSCEPP-----RIRGAGTRELALRAVLYAV-IFLMSVGVGNVLI 76
 QY 61 MIGRYRDRVTRTNLYLGSNAVSDLLILIG-LPFDLYRLWRGRP-----WVFGPLLCR-L 113
 DB 77 VVLGSRRLRTVTNAPFLSLAVSDLLAVACMPFTLL-----PNLMGTPIFGTVVCKAV 130
 QY 114 SLVVGEGCTYATLLHMTALSVERYLAICRPLARVLTTRRRVRLIAVLMAVALLSAGP 173
 DB 131 SYFMGVSVSVST-LSLVAIALERYSAICRPLQARVQWTRSHARVIVATMTMLSGLLWVPY 189
 QY 174 LFLVGVQDQGISVWPGNLGTARIASSPLASSPPLWLSRAPPSPSPGPTAAALFSR 233
 DB 190 PVTAVQ-----PAGPRVLCMRW-- 209
 QY 234 ECRPSPAQLGALRVLMLVTVTAYFFLPFLCLSLYGLIGELW-----SSRRP 280

Db 210 ---PSARVRQTSVLLLL--LFFPVGVMNAVGLISRELYLGLRFDGSDSESQSRVG 264

Qy 281 LRGPAAAGRGHRQTVRLKWSRGSK--DACLOSAPGTAQTGLPLLLAQLWAPLP 338

Db 265 SOGLPGGTGGQPAQNGRCRSTRJAGEDGCGYQLPRSR-----PALEMSALTATPT 319

Qy 339 APPISIPASTRRGGSGIYNLVAL 364

Db 320 GPSGTRPAQAKLLAKKVVRLVI 345

RESULT 12

OPRX_CAVPO STANDARD; PRT; 370 AA.

AC P47748; 1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE Nociceptin receptor (Orphanin FQ receptor) (Kappa-type 3 opioid receptor) (KOR-3) (XOR).

GN OPR1 OR OOR

OS Cavia porcellus (Guinea pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Hystriognathi; Caviidae; Cavia.

OX NCBI_TaxID=10141;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Hartley; TISSUE=Brain;

RA Xie G.;

RL Submitted (DEC-1993) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: RECEPTOR FOR THE NEUROPEPTIDE NOCICEPTIN/ORPHANIN FQ. HAS A POTENTIAL ROLE IN MODULATING A NUMBER OF BRAIN FUNCTIONS, INCLUDING INSTINCTIVE BEHAVIORS AND EMOTIONS. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH INHIBITS ADENYLYL CYCLASE.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; U04369; AAA59332.1; -

DR HSSP; P34996; 1DD0.

DR InterPro; IPR000276; GPCR_Rhodopsn.

DR Pfam; PF00001; 7tm.1.1.

DR PRINTS; PR00237; GPCRHOOPS.

DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; 1.

DR PROSITE; PS0262; G_PROTEIN_RECP_F1_2; 1.

KW G-protein coupled receptor; Transmembrane; Glycoprotein; Phosphorylation; Lipoprotein; Palmitate.

FT DOMAIN 1 50 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 51 77 1 (POTENTIAL).

FT DOMAIN 78 87 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 88 109 2 (POTENTIAL).

FT DOMAIN 110 124 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 125 146 3 (POTENTIAL).

FT DOMAIN 147 165 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 166 188 4 (POTENTIAL).

FT DOMAIN 189 211 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 212 236 5 (POTENTIAL).

FT DOMAIN 237 264 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 265 288 6 (POTENTIAL).

FT DOMAIN 289 300 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 301 322 7 (POTENTIAL).

FT DOMAIN 323 370 CYTOPLASMIC (POTENTIAL).

FT DISULFID 123 200 BY SIMILARITY.

FT LIPID 334 334 PALMITATE (POTENTIAL).

FT CARBOHYD 21 21 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 28 28 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 370 AA; 40789 MW; DAA807CE24283573 CRC64;

Query Match 13.1%; Score 267.5; DB 1; Length 370;

Best Local Similarity 24.9%; Pred. No. 1.4e-11;

Matches 83; Conservative 53; Mismatches 116; Indels 81; Gaps 7;

Qy 17 PWPALPP---CDERRCSPFPGALVPVTAVCICLFVVGSGNVVTVMLIGRYDRMTT 73

Db 27 PNHSGLPPELLLNASHAFPLGLKVTIVGLYLAVCIGGLGNCLVMYVILRHTKMTAT 86

Qy 74 NLYGSMAYSDLLILGLPFDLYRLWRSRPVFGPLLCRLSLYVVGSGCYATVLLHMTALS 133

Db 87 NIYIFNLALADTLVLLTLPFOATDILLGF-WFPGNTLCKTVTAIDYNNFTSTFTTAMS 145

Qy 134 VRYLAICRPLRARVLVTRRRVRLALVLAVALISAGPFLFLVGVQDPGIGSVVPLNG 193

Db 146 VDRYVAICHPIRALDVRTSSKAQVNVAIWALAV-----VGVPVAIMG 189

Qy 194 TARIASSPLASPPPLWLSRAPPPSPSPGPETAEAAALFSRECPSPAQLGALRVMLVVT 253

Db 190 SAQVEDEIEC-----LVIPDPQDYMGVFAVSLPLFS----- 223

Qy 254 AVFFLPFLCLSLYGLIGRELWSSR-----RPLRG----- 283

Db 224 --FIPLVLIISVCYSLMIRRLHGVRLLSGSRKDRNLRLVLLVVAVFVGCWTPVQV 281

Qy 284 ---PAASGRGHRQTVRLKWSRGSKDACL 313

Db 282 FVLVQGLGVQPGSETTVALRFTCALGYVNSCL 314

RESULT 13

NTR2_MOUSE STANDARD; PRT; 417 AA.

AC P70310;

DT 01-NOV-1997 (Rel. 35, Created)

DT 16-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Neurotensin receptor type 2 (NT-R-2) (Low-affinity leucobastine-sensitive neurotensin receptor) (NTRL).

GN NTR2.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BALB/C; TISSUE=Brain;

RX MEDLINE=96388216; PubMed=8795617;

RA Mazella J., Botto J.-M., Guillemare E., Coppola T., Sarret P., Vincent J.-P.;

RT "Structure, functional expression, and cerebral localization of the leucobastine-sensitive neurotensin/neuromedin N receptor from mouse brain."

RT J. Neurosci. 16:5613-5620(1996).

CC -1- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUROTENSIN. IT IS ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- TISSUE SPECIFICITY: EXPRESSED MAXIMALLY IN THE CEREBELLUM, HYPOCAMPUS, PIRIFORM CORTEX AND NEOCORTEX OF ADULT BRAIN.

CC -1- DEVELOPMENTAL STAGE: EXPRESSED POORLY IN 7-DAY-OLD BRAIN. EXPRESSION INCREASES AT DAY 15 TO REACH A MAXIMAL LEVEL IN 35-DAY-OLD BRAIN.

CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS. HIGHEST TO TACHYKININS RECEPTORS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

CC EMBL; U51908; AAB17285.1; -.
CC MGD; MGI:108018; Nter2.
CC InterPro; IPR000276; GPCR_Rhodopsin.
CC InterPro; IPR003986; NT2_rec.
CC InterPro; IPR003984; NT_rec.
CC Pfam; PF00001; 7tm.1; 1.
CC PRINTS; PR00237; GPCR_Rhodopsin.
CC PRINTS; PR01479; NEUROSENSIN.
CC PRINTS; PR01481; NEUROSENSIN2.
CC PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.
CC PROSITE; PS00262; G-PROTEIN_RECEP_F1_2; 1.
CC G-protein coupled receptor; Transmembrane; Lipoprotein; Palmitate.
FT DOMAIN 1 32 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 33 55 1 (POTENTIAL).
FT DOMAIN 56 64 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 65 87 2 (POTENTIAL).
FT DOMAIN 88 109 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 110 131 3 (POTENTIAL).
FT DOMAIN 132 154 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 155 176 4 (POTENTIAL).
FT DOMAIN 177 217 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 218 238 5 (POTENTIAL).
FT DOMAIN 239 298 6 (POTENTIAL).
FT TRANSMEM 299 319 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 320 338 7 (POTENTIAL).
FT TRANSMEM 339 359 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 360 417 BY SIMILARITY.
FT DISULFID 108 194 PALMITATE (POTENTIAL).
FT LIPID 378 378
SQ SEQUENCE 417 AA; 46537 MW; BFFDDBD6507223DD CRC64;

Query Match 13.0%; Score 266; DB 1; Length 417;
Best Local Similarity 32.5%; Pred. No. 2e-11;
Matches 81; Conservative 36; Mismatches 88; Indels 44; Gaps 8;

QY 19 WPALP-----PCDERCSPPFGLALVPTAVCLFVGVGVNVTVMLIGRYDMRT- 71
Db 7 WPPRPSAGLSLEARGVDTLWAKVLFYALVSLIFALGTAGNALSVHVLKARTGPG 66

QY 72 TTNLYLGSMAVDLLILL-GLPFDLYR-LMRSFPWFPGCLRLSLVVGEGCTVATLHM 129
Db 67 RLRYHLSLALSALLLSVPMELYNFVWSHYFWFVGLGCRGYFVRELCAVATVLSV 126

QY 130 TALSVERYLAICPLRARVLVTRRRVALLAVLWAVALLSAGPFLVGVQDPGIVSWP 189
Db 127 ASLSAERCLAVCPPLRARLLTPRTCTRLSLVWVSLGLALPMVIMGQKHE----- 179

QY 190 GLNGTARIASSPLASSPPLMLSRAP-PPSPSGPETAEAAALFSRECRPSAQL---GAL 245
Db 180 -----MERADGPEPAS-----RVCTVLVSRASSRSTFQVKRAGLL 215

QY 246 RVMVLWTTA 254
Db 216 RSLPWLTA 224

RESULT 14
ID SSR5 MOUSE STANDARD; PRT; 362 AA.
AC O08850; O08998;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Somatostatin receptor type 5 (SSR5).
GN SSR5 OR SMSTR5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

RN RP SEQUENCE FROM N.A.
RC STRAIN=129/Svj; TISSUE=Liver;
RX MEDLINE=97444289; PubMed=9300821;
RA Lublin A.L., Diehl N.L., Hochgeschwender U.;
RT "Isolation and characterization of the gene encoding the type 5 mouse
RL (Mus musculus) somatostatin receptor (msst5).";
RN Gene 195:63-66 (1997).
[2]
RN RP SEQUENCE FROM N.A.
RC STRAIN=129/Svj;
RA Moldovan S., Demayo F., Brunicardi P.C.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
[3]
RN RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Liver;
RA Gordon D.F., Woodhense W.W., Wood W.M., Knauf H., James R.A.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
[4]
RN RP SEQUENCE FROM N.A.
RC STRAIN=129/Svj;
RA Baumeister H., Roosterman D., Schafer J., Kreuzer O., Meyerhof W.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: RECEPTOR FOR SOMATOSTATIN-28. THE ACTIVITY OF THIS
CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH INHIBIT ADENYLYL CYCLASE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

Query Match

12.8%; Score 262; DB 1; Length 362;

Best Local Similarity 26.1%; Pred. No. 3.3e-11;
Matches 80; Conservative 45; Mismatches 103; Indels 78; Gaps 10;

QY 5 WNSDGEAREPPWALPCDERRCSPPPLGA---LVPYAVCLCLFVVGSGNVVTVM 61
Db 12 WNSAASSGSHN--WSLVDPVS-----PMGARAVLPVLYLLVC--TVGLGGTLVIY 60

QY 62 LIGRYDMRTTNNLYLGSMAVSDLLILLGLPFDLYRLWRSPWVFGPFLCRLCLSLYVGEGC 121
Db 61 VVLYAKMKTVTVYILNLAVADVFLMLGLFF-LATQNAVSYWPFSGFLCRLVMTLDGIN 119

QY 122 TYATLHMTALSVERYLAICRPLRARVLTTRRRVRAIAVLMAVALISAGPFLFLVGEVQ 181
Db 120 QFTSIFCLMWSVDYLVAVHPLRSARWRPRVAKLASAAVWVFLSLMLPLLVFADVQE 179

QY 182 DPGISVVGLNGTARTASSPLASPPPLWLSRAPPPSPGPTAEAAALFSRCRPSPAQ 241
Db 180 G-----WGTCNLS-----W-----PEPVG 193

QY 242 LGALRVMLWTVTAYFFLPFLCLSLYGLI-----GRELWSSRRPLRGPASGRGRHR 294
Db 194 LWGAFTYTSVLGFGFLVLCVLYLLVVVKUAGMRVGSRR-----RRRSR 243

QY 295 QTVRVL 300
Db 244 KVTMRV 249

RESULT 15
OPRX_PIG STANDARD; PRT; 370 AA.

AC P79232;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Nociceptin receptor (Orphanin FQ receptor) (Kappa-type 3 opioid
DE receptor) (KOR-3) (ORGC) (K3 opiate receptor).
GN ORPL1.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain cortex;
RX MEDLINE=99140612; Pubmed=9988113;
RA Osinski M.A., Pampusch M.S., Murtaugh M.P., Brown D.R.;
RT "Cloning, expression and functional role of a nociceptin/orphanin FQ
RT receptor in the porcine gastrointestinal tract."
RL Eur. J. Pharmacol. 365:281-289(1999).
CC -1- FUNCTION: RECEPTOR FOR THE NEUROPEPTIDE NOCICEPTIN/ORPHANIN FQ.
CC HAS A POTENTIAL ROLE IN MODULATING A NUMBER OF BRAIN FUNCTIONS,
CC INCLUDING INSTINCTIVE BEHAVIORS AND EMOTIONS. THE ACTIVITY OF
CC THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH INHIBITS ADENYLIL
CC CYCLASE (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U72758; AAB39702.1; -.
DR HSSP; P34996; 1DDD.
DR InterPro; IPR00276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODOPS.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.

G-protein coupled receptor; Transmembrane; Glycoprotein;
Phosphorylation; Lipoprotein; Palmitate.

KW DOMAIN 1 50 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 51 77 1 (POTENTIAL).
FT DOMAIN 78 87 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 88 109 2 (POTENTIAL).
FT DOMAIN 110 124 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 125 146 3 (POTENTIAL).
FT DOMAIN 147 165 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 166 188 4 (POTENTIAL).
FT DOMAIN 189 211 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 212 236 5 (POTENTIAL).
FT DOMAIN 237 264 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 265 288 6 (POTENTIAL).
FT DOMAIN 289 300 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 301 322 7 (POTENTIAL).
FT DOMAIN 323 370 CYTOPLASMIC (POTENTIAL).
FT DISULFID 123 200 BY SIMILARITY.
FT LIPID 334 334 PALMITATE (POTENTIAL).
FT CARBOHYD 21 21 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 28 28 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 370 AA; 48E4F3B6DA223BCD CRC64;

Query Match 12.5%; Score 254.5; DB 1; Length 370;
Best Local Similarity 26.1%; Pred. No. 1.1e-10;
Matches 80; Conservative 49; Mismatches 114; Indels 63; Gaps 8;

QY 2 GSPWNGSDGEGAREPPWALPP---CDERRCSPPPLGALVPVTVAVCLCLFVVGSGNVV 58
Db 15 GSPLOQN---LSLLSPNHSLLPPLLNLNASHGAFPLGLKVTIVGLYLAVCVGLLGNCL 71

QY 59 TVMLIGRYDMRTTNNLYLGSMAVSDLLILLGLPF---DLYRLWRSPWVFGPFLCRLSL 115
Db 72 VMXVILRHTKMTATNIYIPNLALADTAVLLTLPFGQTDVLGFG----WPFGNALCKAVI 127

QY 116 YVGEGETYATLHMTALSVERYLAICRPLRARVLTTRRRVRAIAVLMAVALISAGPFLF 175
Db 128 AIDYNNMFTSAFTLTAMSDRYVAICHPIRALDVRTSSKAQAVNVAIWALASI----- 180

QY 176 LVGVEODPGISVVGLNGTARTASSPLASPPPLWLSRAPPPSPGPTAEAAALFSREC 235
Db 181 -----VGVPVAIMGSAQVEDEIEC-----LVEIPAPQDYMGVFAVCIFLFS--- 223

QY 236 RSPQAQLGALRVMLWTVTAYFFLPFLCLSLYGLICRELWSSRRPLRGPASGRGRHRQ 295
Db 224 -----FVIPVLIISVCVSLMVRRLRGVR-----LLSGSREKORN 257

QY 296 TVRVLR 301
Db 258 LRRIIR 263

Search completed: January 1, 2004, 06:23:02
Job time : 14.5439 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 1, 2004, 06:19:59 ; Search time 63.3659 Seconds
(without alignments)
1571.954 Million cell updates/sec

Title: US-09-719-485-5
Perfect score: 2043
Sequence: 1 MGSPWNGSDPGEGAREPPWP.....WQNLHKGKGFADVLLSVL 386

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 23.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriaph.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	673.5	33.0	363	13	O93413 spherooides
2	641.5	31.4	374	13	O93412 spherooides
3	641	31.4	364	11	Q8BWX8 mus musculus
4	607	29.7	366	6	Q8MHZ5 ovis aries
5	435	21.3	559	13	O93414 spherooides
6	356.5	17.4	426	4	Q8NE20 homo sapien
7	354.5	17.4	426	4	Q9HB89 homo sapien
8	339	16.6	145	6	Q97914 ovis aries
9	338	16.5	395	11	Q91276 mus musculus
10	338	16.5	395	11	Q98Z39 mus musculus
11	336	16.4	415	4	Q96AM5 homo sapien
12	335	16.4	415	4	Q9GZ04 homo sapien
13	331.5	16.2	405	11	O55040 mus musculus
14	329	16.1	395	11	Q9JIB1 rattus norv
15	329	16.1	395	11	Q9ESQ4 rattus norv
16	316	15.5	477	5	Q8ITC7 drosophila

17	309	15.1	428	5	Q9VFN4 drosophila
18	309	15.1	430	5	Q8ITC9 drosophila
19	304.5	14.9	412	11	Q9JJI5 rattus norv
20	302.5	14.8	402	11	Q9JIB2 rattus norv
21	285.5	14.0	418	5	O17239 caenorhabdi
22	266.5	13.0	401	13	Q8JF26 xenopus lae
23	264.5	12.9	401	13	Q9DDR0 xenopus lae
24	262.5	12.8	416	11	Q8VIFS mus musculus
25	262.5	12.8	416	11	Q920Q5 mus musculus
26	262	12.8	385	11	Q9JK40 mus musculus
27	261	12.8	595	5	Q8ITD0 drosophila
28	261	12.8	599	5	Q9VFW6 drosophila
29	255.5	12.5	378	5	Q18701 caenorhabdi
30	253.5	12.4	516	4	Q96LC6 homo sapien
31	252.5	12.4	660	5	Q9VFW5 drosophila
32	251.5	12.3	516	4	Q9NVK7 homo sapien
33	250.5	12.3	513	13	Q9DF99 cacoetomus
34	250.5	12.3	658	5	Q8ITD1 drosophila
35	250	12.2	477	13	Q8JID5 carassius a
36	249	12.2	367	6	Q8MI04 ovis aries
37	248	12.1	453	11	Q8BKF6 mus musculus
38	246.5	12.1	314	11	Q8BYG7 mus musculus
39	245.5	12.0	382	11	Q9ERT2 mus musculus
40	241	11.8	370	13	Q8UWLS fugu rubrip
41	240.5	11.8	371	11	Q8BKB0 mus musculus
42	240	11.7	352	6	Q8HZP1 tupia mino
43	240	11.7	450	11	P89005 praomya nat
44	239.5	11.7	517	5	Q8TOS8 drosophila
45	239	11.7	352	11	O88820 rattus norv

ALIGNMENTS

RESULT 1

O93413 PRELIMINARY; PRT; 363 AA.
AC O93413;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Orphan G protein-coupled receptor.
OS Spherooides nephelus.
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Spherooides.
OX NCBI_TaxID=39110;
RN [1]
RP MEDLINE=20092336; PubMed=10628755;
RA Palyha O.C., Feigner S.D., Tan C.P., McKee K.K., Hrenluk D.L.,
RA Gao Y.D., Schlem K.D., Yang L., Morriello G.J., Nargund R.,
RA Patchett A.A., Howard A.D., Smith R.G.;
RT "Ligand activation domain of human orphan growth hormone (GH)
secretagogue receptor (GHS-R) conserved from pufferfish to humans.";
RL Mol. Endocrinol. 14:160-169(2000).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AF082210; AAC33473.1; -
DR InterPro; IPR000276; GPCR_Rhodopn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCR_RHODOPSN.
DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G-PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Receptor; Transmembrane.
SQ SEQUENCE 363 AA; 41324 MW; 74518BD9CD310991 CRC64;

Query Match 33.0%; Score 673.5; DB 13; Length 363;

Best Local Similarity 49.8%; Pred. No. 5.9e-45;

Matches 135; Conservative 39; Mismatches 55; Indels 43; Gaps 4;

OY 31 SPPLGALVPVTAVCCLFVVGSGNVVTVMILIGRDMRTTNNLYLGSMASDLLILG 90

```
Db 31 SLFPASTLIPVTCILIFVVGVTGNTMTIIQYFKDMKTTNLYLSSMAVSDLVFLC 90
Qy 91 LPDLYLWLSRPVWFGPLCRSLVYVGEQCTATLHMTALSVERYLAIICRPLRARVLV 150
Db 91 LPDLYLWLVKVPWLFGEAVCRLYHYIFEGCTSATILHITALSIERLAIISFPLRSKVMV 150
Qy 151 TRRRVRLIALVMAVALLSAGPFLFLVGVQDPCGISVVPGLNGTARIASSPLASSPPLWL 210
Db 151 TRRRVQVILALMCFALVSAPTILFLVGEYD-----NET----- 185
Qy 211 SRAPPPSPGSPETAFAALFSRECRSPAQL--GALRVMLVWTTAYFFLPFLCLSLYLG 268
Db 186 ----HPDNTG-----QCKHTGYAISGQLHIMVSTVYFFCPMLCLLFLYG 229
Qy 269 LIGRELWSSRPLRGPAAASRGHROTQVRL 300
Db 230 SIGCKLWKSNDLQGPCALARSHROTQVRL 261

RESULT 2
ID Q93412 PRELIMINARY; PRT; 374 AA.
AC Q93412;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Orphan G protein-coupled receptor.
OS Spherooides nepheus.
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Spherooides.
OX NCBI_TaxID=39110;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2092336; PubMed=10628755;
RA Palyha O.C., Feighner S.D., Tan C.P., McKee K.K., Hreniuk D.L.,
RA Gao Y.D., Schleim K.D., Yang L., Morriello G.J., Nargund R.,
RA Patchett A.A., Howard A.D., Smith R.G.;
RT "Ligand activation domain of human orphan growth hormone (GH)
RT secretagogue receptor (GHS-R) conserved from pufferfish to humans.";
RL Mol. Endocrinol. 14:160-169(2000).
CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -|- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL: AF082209; AAC33472.1; -.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR InterPro: IPR000524; HTH_GntR.
DR Pfam: PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCRHHODOPS.
DR PROSITE; PS00237; G PROTEIN RECP Fl_1; 1.
DR PROSITE; PS00262; G-PROTEIN RECP_Fl_2; 1.
DR PROSITE; PS00043; HTH_GNTR_FAMILY; 1.
KW G-protein coupled receptor; Receptor; Transmembrane.
SQ SEQUENCE 374 AA; 42324 MW; 2CF9304F004C7A16 CRC64;

Query Match 31.4%; Score 641.5; DB 13; Length 374;
Best Local Similarity 45.8%; Pred. No. 2e-42;
Matches 138; Conservative 45; Mismatches 73; Indels 45; Gaps 8;

Qy 5 WNGS-DGPEGARPPPALPCDERRCSPFLGALVPVAVCLCLFVVGSGNVVTVMLI 63
Db 13 WEGSHNGTAGLELPP-----LNYSIPLAVITVACTVLTGVGVGNVMTLVV 61
Qy 64 GRYDRMRTTNLYLGSMAVSDLLILGLPDLVRLWRSRPWFGPLLCRLSLYVGGCTY 123
Db 62 SYRDMRTTNLYLCSMAVSDLFTVCMPLDLYRMWRYPWRPFGDALCKLFQVSSCTY 121
Qy 124 ATLLHMTALSVERYLAIICRPLRARVLTTRRRVRLIAVMAVALLSNGPFLFVGVQD 183
Db 122 STILICALSVERYLAIICFPLRAKAVTKRRVRLAILLLMTVSLLSAGPFFVMVGVKES 181
Qy 184 GISVFGNGTARIASSPL-ASSPPLWLSRAPPPSPGPTAEAAALFSRECRPS--PA 240
```

```
Db 182 --IMFPN-----SSDLNESSWPL-----EAVDTRECRMTQYAV 212
Qy 241 QGALRVMLVWTTAYFFLPFLCLSLYLYGLIGRELWSSRPLR-GPAASRGHROTQVRY 299
Db 213 ESGLMEAMWLSVFFFPVFCVLTVLYGLIGRRLWLRHRETTINSRVAYRDKSNROTIRM 272
Qy 300 L 300
Db 273 L 273

RESULT 3
ID Q8BWx8 PRELIMINARY; PRT; 364 AA.
AC Q8BWx8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Growth hormone secretagogue receptor type 1 homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Spinal cord;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL: AK049671; BAC33866.1; -.
DR SEQUENCE 364 AA; 40969 MW; 8F1214E58EF3B2E8 CRC64;

Query Match 31.4%; Score 641; DB 11; Length 364;
Best Local Similarity 45.8%; Pred. No. 2.1e-42;
Matches 142; Conservative 41; Mismatches 65; Indels 62; Gaps 10;

Qy 5 WNGSDGPEGAREP-----PWPALPCD---RRRCSPFLGALVPVAVCLCLFVVGSG 55
Db 2 WNAT--PSEPEPNVTLDLDWDASPGNDSDELLPLFPAPLAGVATATCVALFVVGISG 59
Qy 56 NVVTVMICRYDMRTTNLYLGSMAVSDLLILGLPDLVRLWRSRPWFGPLLCRLSL 115
Db 60 NLLTMLVSRFRELRTTNLYLSSMAFSDLLILFCLMPLDLVRLWQYRPNWFGDLCKLFQ 119
Qy 116 YVGEQCTATLHMTALSVERYLAIICRPLRARVLTTRRRVRLIAVMAVALLSAGPFLF 175
Db 120 FVSECTYATVLTITALSVERYFAICFPLRAKVVTKGRVILVILVAVAFCSAGP1FV 179
Qy 176 LVGVEQDPCGISVVPGLNGTARIASSPLASSPPLWLSRAPPPSPGPTAEAAALFSREC 235
Db 180 LVGVEHE-----NGT-----DPRD-----TNEC 197
Qy 236 RS8--PAQIGALRVMLVWTTAYFFLPFLCLSLYLYGLIGRELWSSRPLRGPAAASG--RE 290
Db 198 RATEPAVRSGLTVMVWSSVFFFLPVFCLTVLYSLIGRKLW--RR--RGDAVAGSSLRD 253
Qy 291 RGHROTQVRL 300
Db 254 QNHKQTVKML 263

RESULT 4
ID Q8MHZ5 PRELIMINARY; PRT; 366 AA.
AC Q8MHZ5;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Ghrelin/growth hormone secretagogue receptor.
```

OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1, 2, and 3; TISSUE=pituitary;
RA Murata T., Ikegami R., Morita Y., Shinozaki K.;
RT "Sheep Ghrelin/Growth Hormone Secretagogue Receptor";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY093948; AM19733.1; -
DR EMBL; AY093949; AM19734.1; -
DR EMBL; AY093950; AM19735.1; -
DR InterPro; IPR000237; GPCR Rhodopsin.
DR InterPro; IPR004704; IID_comp.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00613; EIID-AGA; 1.
DR PRINTS; PR00237; GPCR RHODOPSIN.
DR PROSITE; PS00237; G PROTEIN RECEPTOR FL 1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_FL_2; 1.
KW Receptor.
SQ SEQUENCE 366 AA; 41486 MW; 2F276651BC6C5D57 CRC64;

Query Match 29.7%; Score 607; DB 6; Length 366;
Best Local Similarity 42.2%; Pred. No. 9.6e-40;
Matches 129; Conservative 46; Mismatches 79; Indels 52; Gaps 6;

QY 5 WNCSDGPE-----GAREPPHPALPPCD---ERRCSPPGLALVPVTAVALCLFVVGSGN 56
DB 2 WNAATSEELGPNLTLPLDLDAAAPDNDLSLTDELPLPAPLAGVATATCVALFVVGAGN 61

QY 57 VVTVMILIGRYDMRTTNLYLGSMAVSDLLILGLPFDLYRLWSRPVWFGPLLCRLSLY 116
DB 62 LITLVVSRRETRTNNLYLSWAFSDLLIFLCPDLVRLWHYRPWNLGDLCKLFQF 121

QY 117 VEGCTYATLLHMTALSVERLYLAICRPLRARLVTRRRVRLIAVMAVALLSAGPFLFL 176
DB 122 VSESTYASVLTITALSVERLYFAICPLRAKVVITKGRVGLVLAIVAVAFCSAWPIFML 181

QY 177 VGVQDPCISVVPGLNCTARIASPLASSPPLMSRAPPSGPTAEAAALFREC 236
DB 182 VGVHE-----NGT-----DPRD-----TNECR 199

QY 237 PS--PAQALGALVMTWTTAYFFLPFLCLSLYGLIGLWSSRRPLRGPAAASGRGRHR 294
DB 200 ATFAVRSGLTITMVVSSIFFFLPFLVCLTLYSLIGRLKLRWRRRSEVVVGASLRDQNHK 259

QY 295 QTVRVL 300
DB 260 QTVKML 265

RESULT 5
O93414 PRELIMINARY; PRT; 559 AA.
AC O93414;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Orphan G protein-coupled receptor.
OS Spherooides nephelus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Spherooides.
OX NCBI_TaxID=39110;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20092336; PubMed=10628755;
RA Palya O.C., Feighner S.D., Tan C.P., McKee K.K., Hreniuk D.L.,
RA Gao Y.D., Schlem K.D., Yang L., Morriello G.J., Naugund R.,
RA Patchett A.A., Howard A.D., Smith R.G.;

RT "Ligand activation domain of human orphan growth hormone (GH)
secretagogue receptor (GHS-R) conserved from pufferfish to humans."
RL Mol. Endocrinol. 14:160-169(2000).
DR EMBL; AF082211; AAC33474.1; -
DR InterPro; IPR000276; GPCR_Rhodopsin.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCR RHODOPSIN.
DR PROSITE; PS00237; G PROTEIN RECEPTOR FL 1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_FL_2; 1.
KW Receptor.
SQ SEQUENCE 559 AA; 61956 MW; AFB9F35FAF0CD5F6 CRC64;

Query Match 21.3%; Score 435; DB 13; Length 559;
Best Local Similarity 31.2%; Pred. No. 4.5e-26;
Matches 111; Conservative 52; Mismatches 109; Indels 84; Gaps 9;

QY 22 LPPCDERRC---SPPLGALVPVTAVALCLFVVGSGNVVTVMLIGRYDMRTTNLYLG 78
DB 20 LHKCSNQECHWEEPV-FGMIVCVTIYIPLMFLGLNLITLVVWLRPYMRSSLYLIS 78

QY 79 SMAVSDLLILGLPFDLYRLWSRPVWFGPLLCRLSLYVGECTYATLLHMTALSVERLY 138
DB 79 SLAVSDILILLLLPVDLYLWPRWPFGEIECKSTMFSECTFCISILHITFISLERYL 138

QY 139 AICRPLRARLVTRRRVRLIAVMAVALLSAGPFLFLVGE-----QDPGIS----- 186
DB 139 AVCWPPPTAKTVTRRRTRTIIGCIWLGAAISAAPVWVGVBEVQSDQLSGWRESGAW 198

QY 187 -----VBPGL----- 191
DB 199 TKGEKGFIIGERERENDKGLKDBLEMMWKKEMNECGDKNGVTGFKGDKSLEV 258

QY 192 -NGTARIASSPLASSPPLMSRAPPSGPE---TAEAAALFSECRPS---PAQLGA 244
DB 259 GEGTKEQEHGEAAEEAEAKNKKEDGEGGEGDGDGGRWQVDTRECDTYAVSSGL 318

QY 245 LRVLMTWTTAYFFLPFLCLSLYGLIGLWSSRRPLRGPAAASGRGRHRQTVRVL 300
DB 319 LSAMLVLSNMVFLVPLCILGLVGLIGRTLW-----LRS-QISRRDVRNRTVKML 368

RESULT 6
Q8NE20 PRELIMINARY; PRT; 426 AA.
ID Q8NE20
AC Q8NE20;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE G protein-coupled receptor 66.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Strausberg R.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC036543; AAH36543.1; -
DR Genes; HGNC:4518; GPR66.
DR InterPro; IPR000276; GPCR Rhodopsin.
DR InterPro; IPR005390; NeuromedinU.
DR InterPro; IPR005391; NeuromedinU.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCR RHODOPSIN.
DR PROSITE; PS00237; G PROTEIN RECEPTOR FL 1; 1.
DR PROSITE; PS01566; NEUROMEDINUR.
DR PROSITE; PS0262; G_PROTEIN_RECEP_FL_2; 1.
KW Receptor.
SQ SEQUENCE 426 AA; 47322 MW; 7BF8B5DE2E2DC686 CRC64;

Query Match 17.4%; Score 356.5; DB 4; Length 426;

```

Best Local Similarity 35.0%; Pred. No. 4.8e-20;
Matches 98; Conservative 42; Mismatches 81; Indels 59; Gaps 7;

QY 39 VPVAVCLCLFVVGSGNVVTVMLIGRYDRMTTTLNLYGSMVSDLLILL-GLPFDLYR 97
Db 61 MPICATYLLIFVVGAVNGGLTCLVILRHKAMRTPTNYLFSLVSDLLVLLVGLPLEYE 120
QY 98 LMRSPWVFGPLLCRLSLYVGEGCTYATLHMTALSVERVLAICRPLARVLTTRRRVRA 157
Db 121 MHNYPFLVGCGCYFRTLFFEMVCLASVLNVLTALSVERVAVVHPLQARSMTTRAHVR 180
QY 158 LIAVLWAVALLSAGPFLFVGVEODPGISVVPGLNGTARIASSPPLWLSRAPPPS 217
Db 181 VLGAVMGLAMCLSPNTSLHGIRQ---LHV----- 207
QY 218 PPSGPETAAALFSPRECRPSPAQLGALRVMLWTTA--YFFLPFLCLSLYGLIGRELW 275
Db 208 PCRGFVPDSAVCMVLRP-----RALYNNMVQTTALLFFCLPMAIMSVLYLLIGLRLR 259
QY 276 SSRPL-----RGPAAS-----GRERGHRTVRVL 300
Db 260 RERLLMQEAKGRGSAARSRYTCRLQOHDRGRQVTKML 299

RESULT 7
Q9HB89 ID Q9HB89 PRELIMINARY; PRT; 426 AA.
AC Q9HB89; O43664;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE Neuromedin U receptor 1 (Orphan G protein-coupled receptor).
GN NMUR1
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RX MEDLINE=20490668; PubMed=10899166;
RA Raddatz R., Wilson A.E., Artymyshyn R., Bonini J.A., Borowsky B.,
RA Boteju L.W., Zhou S., Kouranova E.V., Nagorny R., Guevarra M.S.,
RA Dai M., Lerman G.S., Vaysse P.J., Brancheck T.A., Gerald C., Forray C.,
RA Adham N.;
RT "Identification and Characterization of Two Neuromedin U Receptors
RT Differentially Expressed in Peripheral Tissues and the Central Nervous
RT System.";
RL J. Biol. Chem. 275:32452-32459(2000).
RN [2]
RX SEQUENCE OF 24-426 FROM N.A.
RX MEDLINE=99000845; PubMed=9782091;
RA Tan C.P., McKee K.K., Liu Q., Palyha O.C., Feighner S.D.,
RA Hreniuk D.L., Smith R.G., Howard A.D.;
RT "Cloning and characterization of a human and murine T-cell orphan G-
RT protein-coupled receptor similar to the growth hormone secretagogue
RT and neurotensin receptors.";
RL Genomics 52:223-229(1998).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL; AF272362; AAC24793.1; -
DR EMBL; AF044601; AAC02680.1; -
DR EMBL; AF044600; AAC02680.1; JOINED.
DR HSP; P02699; I188.
DR InterPro; IPR000276; GPCR_Rhodopsin.
DR InterPro; IPR005390; NeuromedinU.
DR InterPro; IPR005391; NeuromedinU.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCR_Rhodopsin.
DR PRINTS; PR01565; NEUROMEDINUR.
DR PRINTS; PR01566; NEUROMEDINUR.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; 1.
DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL2; 1.
KW G-protein coupled receptor; Receptor; Transmembrane.

SQ SEQUENCE 426 AA; 47350 MW; F8DD781C020F04AA CRC64;
Query Match 17.4%; Score 354.5; DB 4; Length 426;
Best Local Similarity 35.0%; Pred. No. 7e-20;
Matches 98; Conservative 41; Mismatches 82; Indels 59; Gaps 7;

QY 39 VPVAVCLCLFVVGSGNVVTVMLIGRYDRMTTTLNLYGSMVSDLLILL-GLPFDLYR 97
Db 61 MPICATYLLIFVVGAVNGGLTCLVILRHKAMRTPTNYLFSLVSDLLVLLVGLPLEYE 120
QY 98 LMRSPWVFGPLLCRLSLYVGEGCTYATLHMTALSVERVLAICRPLARVLTTRRRVRA 157
Db 121 MHNYPFLVGCGCYFRTLFFEMVCLASVLNVLTALSVERVAVVHPLQARSMTTRAHVR 180
QY 158 LIAVLWAVALLSAGPFLFVGVEODPGISVVPGLNGTARIASSPPLWLSRAPPPS 217
Db 181 VLGAVMGLAMCLSPNTSLHGIRQ---LHV----- 207
QY 218 PPSGPETAAALFSPRECRPSPAQLGALRVMLWTTA--YFFLPFLCLSLYGLIGRELW 275
Db 208 PCRGFVPDSAVCMVLRP-----RALYNNMVQTTALLFFCLPMAIMSVLYLLIGLRLR 259
QY 276 SSRPL-----RGPAAS-----GRERGHRTVRVL 300
Db 260 RERLLMQEAKGRGSAARSRYTCRLQOHDRGRQVTKML 299

RESULT 8
Q97914 ID Q97914 PRELIMINARY; PRT; 145 AA.
AC Q97914;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Growth hormone secretagogues receptor (Fragment).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RX SEQUENCE FROM N.A.
RX TISSUE=Anterior pituitary;
RA Roh S.G., Nie G., Chen C.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AF118636; AAD17249.1; -
DR InterPro; IPR000276; GPCR_Rhodopsin.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCR_Rhodopsin.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; 1.
DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL2; 1.
KW G-protein coupled receptor; Receptor; Transmembrane.
FT NON_TER 1
FT NON_TER 145
FT NON_TER 145
SQ SEQUENCE 145 AA; 16496 MW; F79A7BAB154371EC CRC64;
Query Match 16.6%; Score 339; DB 6; Length 145;
Best Local Similarity 40.1%; Pred. No. 3.7e-19;
Matches 75; Conservative 26; Mismatches 42; Indels 44; Gaps 4;

QY 109 LILCLSLYVGEGCTYATLHMTALSVERVLAICRPLARVLTTRRRVRAIIVLWVALL 168
Db 1 LICKLQFQVSECTYASVLITITALSVERVPAICFPLRAKVVITKGRKVLIAIWAFAFC 60
QY 169 SAGPFLFVGVEODPGISVVPGLNGTARIASSPPLWLSRAPPPSPPGPATAEA 228
Db 61 SAWPFLVGVHE-----NGT-----DPRD----- 81
QY 229 ALFSRECRPS--PAQLGALRVMLWTTA--YFFLPFLCLSLYGLIGRELWSSRRPLGPAA 286
Db 82 ---TNECRATEFAVRSGLLTINWVSSIFFFLPVFCLTVLYSLIGRKLWRRRRSEVVVGA 138

```

1

Query Match 16.4%; Score 336; DB 4; Length 415;
 Best Local Similarity 31.7%; Pred. No. 1.9e-18;
 Matches 79; Conservative 51; Mismatches 73; Indels 46; Gaps 5;

QY 30 CSPPLGALVPVAVCLCLFVGVGNNVTVMLIGRYDRMTTNLYLGSMAVSDLLI-L 88
 DB 37 CGPRSHFFLPVSVVYVPIFVGVIGNVLCVLQHQAMKPTNYLFLSLAVSDLLVLL 96

QY 89 LGLPFDLYRLWRSRPWVFGPCLLCRLSLYVGECHYATLLHMTALSVERYLAICRPLRARV 148
 DB 97 LGMPLVEMWNRNYPFLFGVCGYFKTALFETVCFASILSITTVSERYVAILHPFRKL 156

QY 149 LVTRRRRLALIVLWAVALLSAGPFLFVGVE--QDPGISVVPGLNGTARIASSPLASSP 206
 DB 157 QSTRRRRLALIGIVMGVSLFSLPNTSINGIKFHPFNGSLVPG-----SATCTVIK 208

QY 207 PLWLSRAPPPSPGPTAEAAALFSRECRPSAQLGALRVLMLVTTAYFF-LPFLCLSI 265
 DB 209 PMWI-----YNFIIQVTSFLFVLLPMTVISV 234

QY 266 LYGLIGREL 274
 DB 235 LYVLMALRL 243

RESULT 12

Q9GZQ4 PRELIMINARY; PRT; 415 AA.

ID Q9GZQ4 Q9NR6;
 AC Q9GZQ4; Q9NR6;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Neuromedin U receptor 2 (Neuromedin U receptor-type 2) (G protein-
 coupled receptor TGR-1).
 DE NMUR2 OR NMUR2R OR TGR-1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20490668; PubMed=10899166;
 RA Raddatz R., Wilson A.E., Artymyshyn R., Bonini J.A., Borowsky B.,
 RA Boteju L.W., Zhou S., Kouranova E.V., Nagorny R., Guevarra M.S.,
 RA Dai M., Lerman G.S., Vayssie P.J., Branchek T.A., Gerald C., Forray C.,
 RA Adam N.;
 RT "Identification and Characterization of Two Neuromedin U Receptors
 Differentially Expressed in Peripheral Tissues and the Central Nervous
 System.";
 RL J. Biol. Chem. 275:32452-32459 (2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX PubMed=11010960;
 RA Shan L., Qiao X., Crona J.H., Behan J., Wang S., Laz T., Bayne M.,
 RA Gustafson E.L., Monsma F.J. Jr., Hedrick J.A.;
 RT "Identification of a Novel Neuromedin U Receptor Subtype Expressed in
 the Central Nervous System.";
 RL J. Biol. Chem. 275:39482-39486 (2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Pang L., Wang S., Laz T., Hedrick J.A.;
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 4-415 FROM N.A.
 RX MEDLINE=20351041; PubMed=10894543;
 RA Howard A.D., Wang R., Pong S.-S., Mellin T.N., Strack A., Guan X.-M.,
 RA Zeng Z., Williams D.L., Feighner S.D., Nunes C.N., Murphy B.,
 RA Stair J.N., Yu H., Jiang Q., Clements M.K., Tan C.P., McKee K.K.,
 RA Hreniuk D.L., McDonald T.P., Lynch K.R., Evans J.F., Austin C.F.,
 RA Caskey T., van der Ploeg L.H.T., Liu Q.;
 RT "Identification of receptors for neuromedin U and its role in
 feeding.";

RL Nature 406:70-74 (2000).
 RN [5]
 RP SEQUENCE OF 4-415 FROM N.A.
 RX PubMed=10887190;
 RA Hosoya M., Moriya T., Kawamata Y., Ohkubo S., Fujii R., Matsui H.,
 RA Shintani Y., Fukueumi S., Habata Y., Hinuma S., Onda H., Nishimura O.,
 RA Fujino M.;
 RT "Identification and Functional Characterization of a Novel Subtype of
 Neuromedin U Receptor.";
 RL J. Biol. Chem. 275:29528-29532 (2000).
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 DR ENBL; AF272363; AAG24794.1; -;
 DR ENBL; AF292402; AAG03064.1; -;
 DR ENBL; AF242874; AAF82755.1; -;
 DR ENBL; AB041228; BAB13721.1; -;
 DR InterPro; IPR000276; GPCR Rhodopsin.
 DR InterPro; IPR005390; NeuromedinU.
 DR InterPro; IPR005392; NeuromedinU2.
 DR Pfam; PF00001; 7tm 1; 1.
 DR PRINTS; PR00237; GPCRHOPOPSN.
 DR PRINTS; PR01567; NEUROMEDINUR.
 DR PRINTS; PR01565; NEUROMEDINUR.
 DR PROSITE; PS00237; G PROTEIN RECP F1_1; 1.
 DR PROSITE; PS0262; G PROTEIN RECP F1_2; 1.
 KW G-protein coupled receptor; Receptor; Transmembrane.
 SQ SEQUENCE 415 AA; 47725 MW; C2BACD84B313390F CRC64;

Query Match 16.4%; Score 335; DB 4; Length 415;
 Best Local Similarity 31.7%; Pred. No. 2.3e-18;
 Matches 79; Conservative 51; Mismatches 73; Indels 46; Gaps 5;

QY 30 CSPPLGALVPVAVCLCLFVGVGNNVTVMLIGRYDRMTTNLYLGSMAVSDLLI-L 88
 DB 37 CGPRSHFFLPVSVVYVPIFVGVIGNVLCVLQHQAMKPTNYLFLSLAVSDLLVLL 96

QY 89 LGLPFDLYRLWRSRPWVFGPCLLCRLSLYVGECHYATLLHMTALSVERYLAICRPLRARV 148
 DB 97 LGMPLVEMWNRNYPFLFGVCGYFKTALFETVCFASILSITTVSERYVAILHPFRKL 156

QY 149 LVTRRRRLALIVLWAVALLSAGPFLFVGVE--QDPGISVVPGLNGTARIASSPLASSP 206
 DB 157 QSTRRRRLALIGIVMGVSLFSLPNTSINGIKFHPFNGSLVPG-----SATCTVIK 208

QY 207 PLWLSRAPPPSPGPTAEAAALFSRECRPSAQLGALRVLMLVTTAYFF-LPFLCLSI 265
 DB 209 PMWI-----YNFIIQVTSFLFVLLPMTVISV 234

QY 266 LYGLIGREL 274
 DB 235 LYVLMALRL 243

RESULT 13

O55040 PRELIMINARY; PRT; 405 AA.

ID O55040
 AC O55040;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Orphan G protein-coupled receptor.
 GN GPR66.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA McKee K.K., Tan C.T., Liu J., Palyha O.C., Feighner S.D.,
 RA Hreniuk D.L., Smith R.G., Howard A.D.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

Query Match 16.1%; Score 329; DB 11; Length 395;

Best Local Similarity 30.1%; Pred. No. 6.4e-18;
Matches 75; Conservative 55; Mismatches 73; Indels 46; Gaps 5;

Qy	30	CSPEPLGALYPTAVCLCLFVGVSGNVVTVMLIGRYDMRTTTLNLGSMAYSDLLI-L	88
Db	29	CGPKRSDLSFVSAYALIFLVGVGNLLVCMVIVRHQTLKTPNTNYLFLSLAVSDLLVLL	88
Qy	89	LGLPFDLYRLWRSPWVFGPLLCRLSLYVGEGCTYATLLHMTALSVERYLAICRPLRARV	148
Db	89	LGMPLFIYEMWNNYPFLFGPVGCYFKTALFETVCFASILSVTVTSVERYVAIVHPFRKL	148
Qy	149	LVTRRVRALIAVMAVALLSAGFLELVGV--EQDPGISVWPGNLGTARIASPLASSP	206
Db	149	ESTRRRALRILSLVWSPSVFSLPNTSIHGKFKQHPNGSSVPG-----SATCTVTK	200
Qy	207	PLMLSRAPPPSPSGPETAETAAALFSRECRPSPAQLGALRAVLMWVTYFF-LPFLCLSI	265
Db	201	FMWV-----YNLIIOATSFIFYILPMTLISV	226
Qy	266	LYGLIGREL	274
Db	227	LYYLMGLRL	235

Search completed: January 1, 2004, 06:22:27
Job time : 64.3659 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 1, 2004, 06:19:59 ; Search time 14.4561 Seconds
(without alignments)
1340.261 Million cell updates/sec

Title: US-09-719-485-3
Perfect score: 2155
Sequence: 1 MGSPWNGSDGPGAREPPWP.....DTGGDTVGYTETSANVKTWG 412

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2155	100.0	412	1	MTLR_HUMAN
2	864.5	40.1	366	1	GHRS_PIG
3	863.5	40.1	366	1	GHRS_HUMAN
4	861.5	40.0	364	1	GHRS_RAT
5	614	28.5	257	1	GHRS_MOUSE
6	501	23.2	424	1	NTRI_RAT
7	499.5	23.2	424	1	NTRI_MOUSE
8	497	23.1	418	1	NTRI_HUMAN
9	413	19.2	410	1	NTR2_HUMAN
10	404	18.7	417	1	NTR2_MOUSE
11	401	18.6	416	1	NTR2_RAT
12	370.5	17.2	362	1	SSR5_MOUSE
13	359	16.7	363	1	SSR5_RAT
14	357.5	16.6	418	1	SSR3_HUMAN
15	355.5	16.5	453	1	GP39_HUMAN
16	344	16.0	364	1	SSR5_HUMAN
17	343.5	15.9	380	1	OPRK_CAVPO
18	342.5	15.9	370	1	OPRK_CAVPO
19	339	15.7	380	1	OPRK_RAT
20	335	15.5	380	1	OPRK_HUMAN
21	335	15.5	388	1	SSR4_HUMAN
22	334.5	15.5	367	1	OPRK_MOUSE
23	334.5	15.5	367	1	OPRK_RAT
24	334.5	15.5	370	1	OPRK_PIG
25	334.5	15.5	428	1	SSR3_RAT
26	333.5	15.5	372	1	OPRD_HUMAN
27	332.5	15.4	395	1	TRFR_CHICK
28	332	15.4	372	1	GALS_RAT
29	331.5	15.4	428	1	SSR3_MOUSE
30	330.5	15.3	370	1	OPRK_MOUSE
31	329	15.3	380	1	OPRK_HUMAN
32	327.5	15.2	371	1	GALS_MOUSE
33	327.5	15.2	384	1	SSR4_RAT
					043193 homo sapien
					Q95254 sus scrofa
					Q92847 homo sapien
					O08725 rattus norv
					Q99p50 mus musculus
					P20789 rattus norv
					O88319 mus musculus
					P30989 homo sapien
					O95665 homo sapien
					P70310 mus musculus
					O63384 rattus norv
					O08858 mus musculus
					P30938 rattus norv
					P32745 homo sapien
					O43194 homo sapien
					P35346 homo sapien
					P41144 cavia porce
					P47748 rattus norv
					P34975 cavia porce
					P41145 homo sapien
					P31391 homo sapien
					P35377 mus musculus
					P35370 rattus norv
					P79232 sus scrofa
					P30936 rattus norv
					P41143 homo sapien
					O93603 gallus gall
					O08726 rattus norv
					P30935 mus musculus
					P41146 homo sapien
					P33534 mus musculus
					O88854 mus musculus
					P30937 rattus norv

RESULT 1				
MTLR_HUMAN	15.1	519	1	TLR2_DROME
AC	326	453	1	CCIR_XENLA
DT	325	370	1	GALT_RAT
DT	322.5	384	1	SSR4_MOUSE
DT	322	372	1	OPRD_RAT
DT	321.5	398	1	TRFR_HUMAN
DT	321.5	353	1	APJ_XENLA
DT	320	372	1	OPRD_MOUSE
DT	319.5	387	1	GALS_HUMAN
DT	319.5	398	1	TRFR_BOVIN
DT	318.5	368	1	GALT_HUMAN
DT	318.5	398	1	TRFR_SHEEP
ALIGNMENTS				
PRT; 412 AA.				
SEQUENCE FROM N.A. (ISOFORM A).				
RP	MChee K.K., Tan C.P., Palyha O.C., Liu J., Feighner S.D.,			
RA	Hreniuk D.L., Smith R.G., Howard A.D., van der Ploeg L.H.T.;			
RT	"Cloning and characterization of two human G protein-coupled receptor			
RL	genes (GPR38 and GPR39) related to the growth hormone secretagogue			
RL	and neuropeptide receptors";			
RL	Genomics 46:426-434(1997).			
SEQUENCE FROM N.A. (ISOFORMS A AND B).				
RP	MChee K.K., Tan C.P., Palyha O.C., Hreniuk D.L.,			
RA	Feighner S.D., Tan C.P., McKee K.K., MacNeil D., Cascieri M.A.,			
RA	Pong S.-S., Austin C.P., Figueroa D., Stocco R., Kargman S.,			
RA	Nargund R., Bakshi R., Abramovitz M., Stocco R., Kargman S.,			
RA	O'Neill G., van der Ploeg L.H.T., Evans J., Patchett A.A., Smith R.G.,			
RA	Howard A.D.;			
RT	"Receptor for motilin identified in the human gastrointestinal			
RL	system";			
RL	Science 284:2184-2188(1999).			
SEQUENCE FROM N.A. (ISOFORM A).				
RP	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RL	Wall M.;			
RL	FUNCTION.			
RP	MEDLINE=21219832; PubMed=11322507;			
RA	Smith R.G., Leonard R., Bailey A.R.T., Palyha O.C., Feighner S.D.,			
RA	Tan C.P., McKee K.K., Pong S.-S., Griffin P.R., Howard A.D.;			
RT	"Growth hormone secretagogue receptor family members and ligands.";			
RL	Endocrine 14:9-14(2001).			
CC	-1- FUNCTION: Receptor for motilin.			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	-1- ALTERNATIVE PRODUCTS:			
CC	Event=Alternative splicing; Named isoforms=2;			
CC	Name=A;			
CC	isoId=O43193-1; Sequence=Displayed;			
CC	Name=B;			
CC	isoId=O43193-2; Sequence=VSP_001894;			
CC	-1- TISSUE SPECIFICITY: EXPRESSED ONLY IN THYROID, STOMACH, AND BONE			
CC	MARROW.			
CC	-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.			

P30375 drosophila
P70031 xenopus lae
O88626 rattus norv
P49660 mus musculus
P35333 rattus norv
P34981 homo sapien
P79960 xenopus lae
P32300 mus musculus
O43603 homo sapien
O46639 bos taurus
O60755 homo sapien
Q28596 ovis aries

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AF034632; AAC26081.1; -;
 EMBL; AL137000; CAC19107.1; -;
 Genew; HGNC:4495; GPR38.
 DR MTM; 602885; -;
 DR GO; GO:0005887; C: integral to plasma membrane; TAS.
 DR GO; GO:0007586; P: digestion; TAS.
 DR GO; GO:0007186; P: G-protein coupled receptor protein signalin. . . ; TAS.
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm 1; 1.
 DR PRINTS; PR00237; GPCRRHODOPS.
 DR PROSITE; PS00237; G PROTEIN RECEPTOR FL 1; 1.
 DR PROSITE; PS0262; G PROTEIN RECEPTOR FL 2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW Alternative splicing.
 FT DOMAIN 1 35 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 36 56 1 (POTENTIAL).
 FT DOMAIN 57 74 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 75 94 2 (POTENTIAL).
 FT DOMAIN 95 112 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 113 134 3 (POTENTIAL).
 FT DOMAIN 135 157 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 158 178 4 (POTENTIAL).
 FT DOMAIN 179 246 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 247 270 5 (POTENTIAL).
 FT DOMAIN 271 298 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 299 320 6 (POTENTIAL).
 FT DOMAIN 321 334 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 335 358 7 (POTENTIAL).
 FT DOMAIN 359 412 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 111 235 BY SIMILARITY.
 FT CARBOHYD 192 192 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 301 412 N-LINKED (GLCNAC. . .) (POTENTIAL).
 LVLVLAFLIICWLPFHVGRRIIYINTDSRMWFSQYFNIVALQLFYLASINPILYNLSK 360
 QLFYLSAFINPILYNLSIKRYAAAFKLLARKSRPRGHR
 SRTDAGEVAGDTGDTGVGTETGANKVTMG -> RKWSRRG
 SKDACLQSPGTAQTLGPIPLLAQLWAPLPAPFPISIPAS
 TRGGSGGIYLLVALPRWQVHLKGRFADVLLSVL
 (in isoform B).
 /FTID=VSP_001894.

SQ SEQUENCE 412 AA; 45344 MW; C13FF6165012DEF3 CRC64;
 Query Match 100.0%; Score 2155; DB 1; Length 412;
 Best Local Similarity 100.0%; Pred. No. 66-137;
 Matches 412; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGSPWNGSDGEGAREPPALPCDERRCSPPFLGALVPTAVCLFLVVGSGNVTV 60
 DB 1 MGSPWNGSDGEGAREPPALPCDERRCSPPFLGALVPTAVCLFLVVGSGNVTV 60
 QY 61 MLIGRYDMRTTNLYLGSAVSDLLILLGLPDLVRLMSRPWFGPLLCRLSLVYVGE 120
 DB 61 MLIGRYDMRTTNLYLGSAVSDLLILLGLPDLVRLMSRPWFGPLLCRLSLVYVGE 120
 QY 121 CTYATLLHMTALSVERVLAICRPLARVLVTRVRVLAIVLWALLSAGPFLVGVGE 180
 DB 121 CTYATLLHMTALSVERVLAICRPLARVLVTRVRVLAIVLWALLSAGPFLVGVGE 180
 QY 181 QDGISVVPGLNGTARIASSPLWLSRAPPPSPGSPETAATAAALFSRECRPSPA 240
 DB 181 QDGISVVPGLNGTARIASSPLWLSRAPPPSPGSPETAATAAALFSRECRPSPA 240
 QY 241 QLGAALRVMLVTTAYFPFLPCLSLYGLIGRLWSRRPLRGPAAASGRGRHQRVRL 300
 DB 241 QLGAALRVMLVTTAYFPFLPCLSLYGLIGRLWSRRPLRGPAAASGRGRHQRVRL 300

QY 301 LVLVLAFLIICWLPFHVGRRIIYINTDSRMWFSQYFNIVALQLFYLASINPILYNLSK 360
 DB 301 LVLVLAFLIICWLPFHVGRRIIYINTDSRMWFSQYFNIVALQLFYLASINPILYNLSK 360
 QY 361 KYRAAFAKLLARKSRPRGHRSDRTAGEVAGDTGDTGVGTETGANKVTMG 412
 DB 361 KYRAAFAKLLARKSRPRGHRSDRTAGEVAGDTGDTGVGTETGANKVTMG 412

RESULT 2

GHSR_PIG ID GHSR_PIG STANDARD; PRT; 366 AA.
 AC Q95254; Q95255;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Growth hormone secretagogue receptor type 1 (GHS-R) (GH-releasing peptide receptor) (GHRP) (Ghrelin receptor).
 GN GHSR.
 OS Sus acrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RC SEQUENCE FROM N.A. (ISOFORMS 1A AND 1B).
 RC STRAIN=Yorkshire; TISSUE=Pituitary;
 RX MEDLINE=96337998; PubMed=8688086;
 RA Howard A.D., Feighner S.D., Cully D.F., Arena J.P., Hreniuk D.L., Liberator P.A., Rosenblum C.I., Hamelin M., Hreniuk D.L., Palyha O.C., Anderson J., Paress P.S., Diaz C., Chou M., Liu K.K., McKee K.K., Pong S.-S., Chaung L.-Y., Elbrecht A., Dashkevicz M., Heavens R., Rigby M., Sirinathsinghji D.J.S., Dean D.C., Melillo D.G., Patchett A.A., Nargund R., Griffin P.R., Demartino J.A., Gupta S.K., Schaeffer J.W., Smith R.G., van der Ploeg L.H.T.;
 RA "A receptor in pituitary and hypothalamus that functions in growth hormone release.";
 RT Science 273:974-977(1996).
 RL CC -I- FUNCTION: Receptor for ghrelin, coupled to G-alpha-11 proteins. Stimulates growth hormone secretion. Binds also other growth hormone releasing peptides (GHRP) (e.g. Met-enkephalin and GHRP-6) as well as non-peptide, low molecular weight secretagogues (e.g. L-692,429, MK-0677, adenosine).
 CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -I- ALTERNATIVE PRODUCTS.
 CC Event-Alternative splicing; Named isoforms=2;
 CC Name=1A;
 CC IsoId=Q95254-1; Sequence=Displayed;
 CC Name=1B;
 CC IsoId=Q95254-2; Sequence=VSP_001918, VSP_001919;
 CC -I- TISSUE SPECIFICITY: PITUITARY AND HYPOTHALAMUS.
 CC -I- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; U60178; AAC48630.1; -;
 EMBL; U60180; AAC48631.1; -;
 InterPro; IPR000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm 1; 1.
 DR PRINTS; PR00237; GPCRRHODOPS.
 DR PROSITE; PS00237; G PROTEIN RECEPTOR FL 1; 1.
 DR PROSITE; PS0262; G PROTEIN RECEPTOR FL 2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW Alternative splicing.
 FT DOMAIN 1 40 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 41 66 1 (POTENTIAL).
 FT DOMAIN 67 72 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 73 96 2 (POTENTIAL).

```

FT DOMAIN 97 117 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 118 139 3 (POTENTIAL).
FT DOMAIN 140 162 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 163 183 4 (POTENTIAL).
FT DOMAIN 184 211 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 212 235 5 (POTENTIAL).
FT DOMAIN 236 263 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 264 285 6 (POTENTIAL).
FT DOMAIN 286 302 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 303 326 7 (POTENTIAL).
FT DOMAIN 327 366 CYTOPLASMIC (POTENTIAL).
FT DISULFID 116 198 BY SIMILARITY.
FT CARBOHYD 13 13 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 266 289 AVVVFAPILCWLPFHVGRYLFSSK -> GGSQCALELSLPG
PLHSSCLFSSP (in isoform 1B).
/FTID=VSP_001918.
FT VARSPLIC 290 366 Missing (in isoform 1B).
/FTID=VSP_001919.
SQ SEQUENCE 366 AA; 41194 MW; 2C850B3BF61B7C1C CRC64;
Query Match 40.1%; Score 864.5; DB 1; Length 366;
Best Local Similarity 45.1%; Pred. No. 5.9e-51;
Matches 189; Conservative 55; Mismatches 104; Indels 71; Gaps 10;
QY 5 WNGSDGEPA-----REPPWALPPCD---ERRCSPPLGALVPVTVAVCLCLFVWVSGN 56
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2 WNATPSEPGPNLTLPDLGWDAPPENDSLVEELLPLFPPTLLAGVATATCVALFVWGIAGN 61
QY 57 VVTVMILGRVDRMTTNLYLGSMAVSDILILGLPDLVRLWRSRPWVGPPLICRLSLY 116
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
62 LLTMLVRSRPREMTTNLYLGSMAVSDILILGLPDLVRLWRSRPWVGPPLICRLSLY 121
QY 117 VGGCTVATLHTLTALSVRYLAICRLRLRVLRVTRRRVRLAVLWALLSAGPPLFL 176
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
122 VSSECTVATLHTLTALSVRYLAICRLRLRVLRVTRRRVRLAVLWALLSAGPPLFL 181
QY 177 VGVQDQGISVWPLGNTARIASPLASPPPLMSRAPPPSPGPETAABAAAFSRECR 236
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
182 VGVQHD-----NCT-----DPRD-----TNECR 199
QY 237 PS--PAQLGALRWMLWTTAYFPLPFLCLILYGLIGRELWSSRRPLRGPAAAG---RER 291
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
200 ATEFAVRSGLTVMWVSSVFFFLPFLVCLVLYSLIGRLKLV---RRKRGEAAVGSLLRQ 256
QY 292 GHRQTVRLVAVLARIICLWLPFHVGRIV--INTEDSRMVFQYFNIVLQFLVLSA 348
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
257 NHQTVRLVAVLARIICLWLPFHVGRIV--INTEDSRMVFQYFNIVLQFLVLSA 316
QY 349 SINPILNLSKKYRAAFAKLLARKSRPRGFRHSRDTAGEVAGDTGGDTVGVYTETSAN 407
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
317 AINPILNLSKKYRAVAVKLLGPEPFSQKLSLTKDESSR-----ANTESSIN 365

```

RESULT 3

```

GHSR_HUMAN
ID GHSR_HUMAN STANDARD; PRT; 366 AA.
AC Q92847; Q92848; Q96R37;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Growth hormone secretagogue receptor type 1 (GHS-R) (GH-releasing
DE peptide receptor) (GHRP) (GHRP) (GHRP) (GHRP) (GHRP) (GHRP) (GHRP)
GN GHSR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1A AND 1B).
RC TISSUE=Plutary; PubMed=8688086;
RX MEDLINE=96337998;
RA Howard A.D., Feighner S.D., Cully D.F., Arena J.P.,

```

```

RA Liberato P.A., Rosenblum C.I., Hamelin M., Hreniuk D.L.,
RA Palyha O.C., Anderson J., Pares P.S., Diaz C., Chou M., Liu K.K.,
RA McKee K.K., Pong S.-S., Chung L.-Y., Elbrecht A., Dashkevich M.,
RA Heavens R., Rigby M., Sirinathsinghji D.J.S., Dean D.C., Melillo D.G.,
RA Patchett A.A., Nargund R., Griffin P.R., Demartino J.A., Gupta S.K.,
RA Schaeffer J.M., Smith R.G., van der Ploeg L.H.T.,
RT "A receptor in pituitary and hypothalamus that functions in growth
RT hormone release."
RL Science 273:974-977(1996).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1A AND 1B).
RX MEDLINE=21255649; PubMed=11356716;
RA Petersen S., Rasch A.C., Penschorn M., Bell F.U., Schulte H.M.;
RT "Genomic structure and transcriptional regulation of the human growth
RT hormone secretagogue receptor."
RL Endocrinology 142:2649-2659(2001).
RN [3]
RP FUNCTION.
RX MEDLINE=21219832; PubMed=11322507;
RA Smith R.G., Leonard R., Bailey A.R.T., Palyha O.C., Feighner S.D.,
RA Tan C.P., McKee K.K., Pong S.-S., Griffin P.R., Howard A.D.;
RT "Growth hormone secretagogue receptor family members and ligands."
RL Endocrine 14:9-14(2001).
RN [4]
RP FUNCTION.
RX MEDLINE=20087959; PubMed=10604470;
RA Kojima M., Hosoda H., Date Y., Nakazato M., Matsuo H., Kangawa K.;
RT "Ghrelin is a growth-hormone-releasing acylated peptide from
RT stomach."
RL Nature 402:656-660(1999).
CC -1- FUNCTION: Receptor for ghrelin, coupled to G-alpha-11 proteins.
CC Stimulates growth hormone secretion. Binds also other growth
CC hormone releasing peptides (GHRP) (e.g. Met-enkephalin and GHRP-6)
CC as well as non-peptide, low molecular weight secretagogues (e.g.
CC L-692,429, MK-0677, adenosine).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1A;
CC IsoId=Q92847-1; Sequence=Displayed;
CC Name=1B;
CC IsoId=Q92847-2; Sequence=VSP_001916, VSP_001917;
CC -1- TISSUE SPECIFICITY: PITUITARY AND HYPOTHALAMUS.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
```

```

FT DOMAIN 97 117 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 118 139 3 (POTENTIAL).
FT DOMAIN 140 162 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 163 183 4 (POTENTIAL).
FT DOMAIN 184 211 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 212 235 5 (POTENTIAL).
FT DOMAIN 236 263 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 264 285 6 (POTENTIAL).
FT DOMAIN 286 302 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 303 326 7 (POTENTIAL).
FT DOMAIN 327 366 CYTOPLASMIC (POTENTIAL).
FT DISULFID 116 198 BY SIMILARITY.
FT CARBOHYD 13 13 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT 266 AVVFAFICLWLPFHVGRYLSKLS -> GGSQALRLSLAG
PILSLCLPLSL (in isoform 1B).
/FTID=VSP_001916.
/FTID=VSP_001917.
FT VARSPLIC 290 366 Missing (in isoform 1B).
SQ SEQUENCE 366 AA; 41328 MW; D1B62710DA9DC0C6 CRC64;

Query Match 40.18; Score 863.5; DB 1; Length 366;
Best Local Similarity 44.78; Pred. No. 6.9e-51;
Matches 186; Conservative 59; Mismatches 106; Indels 65; Gaps 9;

QY 5 WNGSDGPEGA-----REPPWALPPCD---ERRCSPPFLGALVPVTVAVCLCLFVVGSGN 56
DB 2 WNATPSEPGFNLTLDLDWDASPGNDSLGDELQLQLEPAPLLAGVATCTVALFVVGAGN 61

QY 57 VVTVMILGRYDRMRTTNNLYLGSMAVSDLLILGLPDLVRLWRSRPWVFGPLLCRLSLY 116
DB 62 LLTGLVTVSRPRELRTTNNLYLGSMAVSDLLILGLPDLVRLWRSRPWVFGPLLCRLSLY 121

QY 117 VGSCTATLHTALSVRYLAICRPLBARVLVTRRRVATLAVLWALLSAGFLFL 176
DB 122 VBSCTATLHTALSVRYLAICRPLBARVLVTRRRVATLAVLWALLSAGFLFL 181

QY 177 VGYEODPGISVWPGCLNTARIASSPLASPLMLSRPPSPGPTAEAAALFSGREC 236
DB 182 VGYEHE-----NGT-----DP-W-----DTNECR 199

QY 237 PS--PAQGLALRVMLVTTAYFPFLCLSLYLGLIGRELWSSRRPLRGPAASGRGRGHR 294
DB 200 PTEFAVRSGLTVMWVSSIFFFLPVFLVCLVLSLIGRLWRRRGDAVVGASLRDQNHK 259

QY 295 QTVRVLLVVLVLAFLICWLPFHVGRYIYINTEDS---RMVYFQYFNIVALQLPYLSASIN 351
DB 260 QTVKMLAVVVFATILCWLPHVGRYLSFKSFEPGSLKIAQISQYCNLVSFLVFLYLSAAN 319

QY 352 PILYNLISKYRAAFAKLLARKSRPRGFRHSRDGTAGEVAGDTGDTVGYTETSAN 407
DB 320 PILYNLISKYRVAVFLLGFEFPFSQRLKSLDKDESSR-----AWTESSIN 365

RESULT 4
GHSR_RAT STANDARD; PRT; 364 AA.
AC Q08725;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Growth hormone secretagogue receptor type 1 (GHS-R) (GH-releasing
DE peptide receptor) (GHRP) (Ghrelin receptor).
GN GHSR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary;
RX MEDLINE=97246555; PubMed=9092793;
RA McKee K.K., Palyha O.C., Feighner S.D., Hreniuk D.L., Tan C.P.,
Phillips M.S., Smith R.G., der Ploeg L.H.T., Howard A.D.;
"Molecular analysis of rat pituitary and hypothalamic growth hormone
secretagogue receptors.";
Mol. Endocrinol. 11:415-423(1997).
[2]
SEQUENCE OF 1-240 FROM N.A.
RC STRAIN=Wistar; TISSUE=Pituitary;
RX MEDLINE=98100386; PubMed=9437732;
RA Yokote R., Sato M., Matsubara S., Ohye H., Niimi M., Murao K.,
Takahara J.;
"Molecular cloning and gene expression of growth hormone-releasing
peptide receptor in rat tissues.";
Peptides 19:15-20(1998).
[3]
FUNCTION
RX MEDLINE=20067959; PubMed=10604470;
RA Kojima M., Hosoda H., Date Y., Nakazato M., Matsuo H., Kangawa K.;
"Ghrelin is a growth-hormone-releasing acylated peptide from
Stomach.";
Nature 402:656-660(1999).
RT Nature 402:656-660(1999).
CC -I- FUNCTION: Receptor for ghrelin, coupled to G-alpha-11 proteins.
CC Stimulates growth hormone secretion. Binds also other growth
CC hormone releasing peptides (GHRP) (e.g. Met-enkephalin and GHRP-6)
CC as well as non-peptide, low molecular weight secretagogues (e.g.
CC L-692,429, MK-0677, adenosine) (By similarity).
CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
CC -I- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL: U94321; AAC53156.1; -
DB EMBL: AB001982; BAA21777.1; ALT INIT.
DB InterPro: IPR000276; GPCR_Rhodopsn.
DB Pfam: PF00001; 7tm1; 1.
DB PRINTS: PRO0237; GPCRHHODOPSN.
DB PROSITE: PS00237; G-PROTEIN RECEPTOR FL1; 1.
DB PROSITE: PS00262; G-PROTEIN RECEPTOR FL2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 40 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 41 66 1 (POTENTIAL).
FT DOMAIN 67 72 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 73 96 2 (POTENTIAL).
FT DOMAIN 97 117 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 118 139 3 (POTENTIAL).
FT DOMAIN 140 182 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 163 183 4 (POTENTIAL).
FT DOMAIN 184 211 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 212 235 5 (POTENTIAL).
FT DOMAIN 236 263 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 264 285 6 (POTENTIAL).
FT DOMAIN 286 302 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 303 326 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 327 364 BY SIMILARITY.
FT DISULFID 115 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 13 13 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 26 26 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 364 AA; 40963 MW; DCF559BE061EE9 CRC64;

Query Match 40.0%; Score 861.5; DB 1; Length 364;
Best Local Similarity 48.3%; Pred. No. 9.3e-51;
Matches 185; Conservative 52; Mismatches 81; Indels 65; Gaps 11;

QY 5 WNGSDGPEGAPEP-----PWPALPPCD---ERRCSPPFLGALVPVTVAVCLCLFVVGSG 55
DB 2 WNAT--PSEEPFNVTLDLDWDASPGNDSLGDELQLQLEPAPLLAGVATCTVALFVVGISG 59

QY 56 NVVTVMILGRYDRMRTTNNLYLGSMAVSDLLILGLPDLVRLWRSRPWVFGPLLCRLSL 115
```

```

Db 60 NLLTMLVSRPRELRTTNLYLSMAFSDLLIFCLMPLDLVRLWQYRPMWFGDLLCKLFQ 119
Qy 116 YVGGCTYATLLHMTALSVERYLAICRPLARVLVTRRRVRLALIAVLWAVALLSAGPFLF 175
Db 120 FVSECTYATVLTITALSVERYPAICPLRAKVVTKGRVKLVILVIAVAFCSAGPIFV 179
Qy 176 LVGVEQDGISSVVGCLNGTARIASSPLASSPPLWLSRAPPPSPGPGTETAAALPFRRC 235
Db 180 LVGVEH-----NGT-----DPRD-----TNEC 197
Qy 236 RPS--PAQLGALRVMLVMTAYTFELPFLCLSLYGLIGRELWSSRRPLRGPASG---RE 290
Db 198 RATEFAVRSGLLTVMWVSSVFFFLPVCLTVLSLIGRLW--RR--RGDAVAGSLRD 253
Qy 291 RGRHQTVRLLVVLVLAIFICWLPVHVGRIIYINTEDS---RMYFSQYFNIVALQLFYL 347
Db 254 QNHKQTVKMLAVVFAILCWLPHVGRYLFKSFEPGSLIAQISQYCNLVSPVLYLS 313
Qy 348 ASINPILYNISKYRAAAFKLL 370
Db 314 AAINPILYNIMSKYRVAVFKLL 336

```

RESULT 5

```

ID GHRS MOUSE STANDARD; PRT; 257 AA.
AC Q99P50; Q91282;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Growth hormone secretagogue receptor type 1 (GHS-R) (GH-releasing
  peptide receptor) (GHRP) (Ghrelin receptor) (Fragment).
GN GHRS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE OF 1-183 FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Hypothalamus;
RA Kacsob B.;
RT "Cloning of mouse ghrelin/growth hormone secretagogue receptor cDNA by
  rapid amplification of cDNA ends (RACE).";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
SEQUENCE OF 73-257 FROM N.A.
RP STRAIN=129S3/SvimJ;
RA Peng X., Frohman L.A., Kineman R.D.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Receptor for ghrelin, coupled to G-alpha-11 proteins.
CC Stimulates growth hormone secretion. Binds also other growth
CC hormone releasing peptides (GHRP) (e.g. Met-enkephalin and GHRP-6)
CC as well as non-peptide, low molecular weight secretagogues (e.g.
CC L-692,429, MK-0677, adenosine) (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AY056474; AAL13336.1; -
CC EMBL; AF332997; AAG61141.1; -
CC InterPro; IPR000276; GPCR_Rhodpsn.
CC Pfam; PF00001; 7tm.1; 1.
CC PRINTS; PR00237; GPCRHHODPSN.
CC PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
CC PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.

```

```

KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 35
FT TRANSMEM 36 58
FT DOMAIN 59 77
FT TRANSMEM 78 100
FT DOMAIN 101 125
FT TRANSMEM 126 148
FT DOMAIN 149 160
FT TRANSMEM 161 183
FT DOMAIN 184 211
FT TRANSMEM 212 234
FT TRANSMEM 235 257
FT DISULFID 115 197
FT CARBOHYD 13 13
FT CARBOHYD 26 26
FT CARBOHYD 187 187
FT NON TER 257 257
SQ SEQUENCE 257 AA; 28743 MW; D6PA21EAB0E30791 CRC64;

```

Query Match 28.5%; Score 614; DB 1; Length 257;
 Best Local Similarity 45.1%; Pred. No. 2.1e-34;
 Matches 137; Conservative 39; Mismatches 66; Indels 62; Gaps 10;

```

Qy 5 WNGSDGPEGAREP-----PWPALPDCD-----ERRCSFFPLGALVPVTAVALCLFVWCVSG 55
Db 2 WNAT--PSEPEPNVTLDLWDASPGNLSLDELPLFPAPPLAGVTATCAVLFVWGISS 59
Qy 56 NVVTMLIGRYDRMTTNLYLGSMASVSDLLILGLPFDLYLRWRSPWVFGPLLCRLSL 115
Db 60 NLUTMLVSRPRELRTTNLYLSMAFSDLLIFCLMPLDLVRLWQYRPMWFGDLLCKLFQ 119
Qy 116 YVGGCTYATLLHMTALSVERYLAICRPLARVLVTRRRVRLALIAVLWAVALLSAGPFLF 175
Db 120 FVSECTYATVLTITALSVERYPAICPLRAKVVTKGRVKLVILVIAVAFCSAGPIFV 179
Qy 176 LVGVEQDGISSVVGCLNGTARIASSPLASSPPLWLSRAPPPSPGPGTETAAALPFRRC 235
Db 180 LVGVEH-----NGT-----DPRD-----TNEC 197
Qy 236 RPS--PAQLGALRVMLVMTAYTFELPFLCLSLYGLIGRELWSSRRPLRGPASG---RE 290
Db 198 RATEFAVRSGLLTVMWVSSVFFFLPVCLTVLSLIGRLW--RR--RGDAVAGSLRD 253
Qy 291 RGRH 294
Db 254 QNHK 257

```

RESULT 6

```

ID NTR1 RAT STANDARD; PRT; 424 AA.
AC P20789;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Neurotensin receptor type 1 (NT-R-1) (High-affinity levocabastine-
  insensitive neurotensin receptor) (NTRH).
GN NTSR1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=90297956; PubMed=1694443;
RA Tanaka K., Masu M., Nakanishi S.;
RT "Structure and functional expression of the cloned rat neurotensin
  receptor.";
RL Neuron 4:847-854(1990).
CC -1- FUNCTION: RECEPTOR FOR THE TRIPEPTIDE NEUROTENSIN. IT IS
CC ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
CC CALCIUM SECOND MESSENGER SYSTEM.

```

CC	-1-	SUBCELLULAR LOCATION: Integral membrane protein.
CC	-1-	SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC	HIGHEST TO TACHYKININS RECEPTORS.	
DR	P1R; JH0164; JH0164.	
DR	InterPro; IPR000276; GPCR_Rhodopsn.	
DR	InterPro; IPR003985; NTL_Fec.	
DR	InterPro; IPR003984; NT_Fec.	
DR	Pfam; PF00001; 7tm 1; 1.	
DR	PRINTS; PR00237; GPCRRHODOPSN.	
DR	PRINTS; PR01479; NEUROTENSINR.	
DR	PRINTS; PR01480; NEUROTENSINR.	
DR	PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.	
DR	PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.	
KW	G-protein coupled receptor; Transmembrane; Glycoprotein;	
KW	Phosphorylation; Lipoprotein; Palmitate.	
FT	DOMAIN 1	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 65	1 (POTENTIAL).
FT	DOMAIN 88	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 97	2 (POTENTIAL).
FT	DOMAIN 122	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 144	3 (POTENTIAL).
FT	DOMAIN 166	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 189	4 (POTENTIAL).
FT	DOMAIN 211	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 236	5 (POTENTIAL).
FT	DOMAIN 261	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 309	6 (POTENTIAL).
FT	DOMAIN 331	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 349	7 (POTENTIAL).
FT	DOMAIN 373	CYTOPLASMIC (POTENTIAL).
FT	CARBOHYD 4	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 38	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 42	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID 142	BY SIMILARITY.
FT	LIPID 388	PALMITATE (POTENTIAL).
SQ	SEQUENCE 424 AA; 47054 MW; A9C2F7EAF8D9BCD3 CRC64;	
Query Match		
Best Local Similarity 32.2%; Score 501; DB 1; Length 424;		
Matches 125; Conservative 59; Mismatches 110; Indels 86; Gaps 11;		
QY	39 VPTAVCLCLFFVGVSGNVVTVMILGR---	YRD MRTTNLYLGNAVSDLLI-LLGLPPD 94
DB	65 VLVTAIYALFVVGTVGNSVTAFTLARKKSLSQSTVHVHLGSLSDLILLAMPVE 124	
QY	95 LYSRLNRSRPWFGPLLCSLLYVGECTVATLHMNTLSVERYLACRPLRARVLVTR 153	
DB	125 LYNFIVHHHPWAEGACRGYIFURDACTATANLVASLSVERYLAICHPFKATLMRSR 184	
QY	154 RVRALIAVLMAVALLSAGPFLFVLGVBDPGISVVVPGINGTARTASSGPLLWSRA 213	
DB	185 RTKKFSIAIWLASALLAIFMLTWGLQRSGDTHFG-----GLVCTPIVDT----- 231	
QY	214 PPPSPSPGPETAABAAALFSRECRPSAQLGALRWMLVTTAYFFL-PFLCLSIYLGIGR 272	
DB	232 -----ATVKVVIQVNTFMSEFLFPMVISILENTVIAN 262	
QY	273 ELWSSRRPLRCGPAAAGR-----ERGHQTVR-----VLLVVVLAFII 309	
DB	263 KLTWM---VHQAAEQGGVCVTGTHTNGLEHSTFNMTIEPFRQAULRHGVLVLRVAVIAFW 319	
QY	310 CWLPFHVGRITTYINTDSR---MMVFQSFYNIVALQLFYLSASINPILYNLSISKYRAA 365	
DB	320 CWLPYHVRRLMFCYISDEQWTTFELDFYHFYFMLTNALFVSSAINPILYNLVSANFRQV 379	
QY	366 AFKLLLA-----RKSRP 377	
DB	380 FLSTACLCPGWRHRXKRKP 399	
RESULT 7		
ID	NTRI_MOUSE	STANDARD; PRT; 424 AA.
ID	NTRI_MOUSE	

```

AC 088319;
AD 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurotensin receptor type 1 (NT-R-1).
DN NTSR1 OR NTSR (Mouse).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RC Snider J., Sano H., Ohta M.;
RT "Neurotensin receptor type 1";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUROTENSIN. IT IS
CC ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
CC CALCIUM SECOND MESSENGER SYSTEM.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC HIGHEST TO TACHYKININS RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB017027; BAA33013.1; -.
CC MGD; MGI:97386; Ntsr.
CC InterPro; IPR000276; GPCR Rhodpsn.
CC InterPro; IPR003985; NTL_rec.
CC InterPro; IPR003984; NT_rec.
CC Pfam; PF00001; 7tm1; 1.
CC PRINTS; PR00237; GPCRRHODOPSN.
CC PRINTS; PR01479; NEUROTHODOPSN.
CC PRINTS; PR01480; NEUROTSNR.
CC PROSITE; PS00337; G_PROTEIN_RECEP_F1_1; 1.
CC PROSITE; PS02662; G_PROTEIN_RECEP_F1_2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein;
CC KW Phosphorylation; Lipoprotein; Palmitate.
CC DOMAIN 1 64
CC FT TRANSMEM 65 87
CC FT DOMAIN 88 96
CC FT TRANSMEM 97 121
CC FT DOMAIN 122 143
CC FT TRANSMEM 144 165
CC FT DOMAIN 166 188
CC FT TRANSMEM 189 210
CC FT DOMAIN 211 235
CC FT TRANSMEM 236 260
CC FT DOMAIN 261 308
CC FT TRANSMEM 309 330
CC FT DOMAIN 331 348
CC FT TRANSMEM 349 372
CC FT DOMAIN 373 424
CC FT CARBOHYD 4 4
CC FT CARBOHYD 38 38
CC FT CARBOHYD 42 42
CC FT CARBOHYD 211 211
CC FT DISULFID 141 224
CC FT LIPID 388 388
CC SEQUENCE 424 AA; 47216 MW; 8E9A723171A48711 CRC64;
CC -----
Query Match 23.2%; Score 499.5; DB 1; Length 424;
Best Local Similarity 33.3%; Pred. No. 1.5e-26;
Matches 126; Conservative 57; Mismatches 114; Indels 81; Gaps 11;
QY 39 VPVTAVCLCLFVGVGSGNVVTVMLIGR---YRDMRTTNTLYLGSMVSDLLI-LLGLPFD 94

```


Db 64 VLTAVYALFVVGTVGNVTYTAFTLARKKSLQSLQSTVHYHLSLALSDLLTLLAMPVE 123
QY 95 LYR-LWRSRPWFVFLCRLSLVYEGCTVATLHMTALSVERYLAICRPLARVLVTRR 153
Db 124 LYNFVWHFPAFGDAGCRGYFYRDACTATALNVASLVERYLAICHPFKATLMSRS 193
QY 154 RVRLIAVLAVALLSAGPFLVGVQDGPISVVPGLNGTARIASSPLSLMSRA 213
Db 184 RTKFGISAIWLASALLVPMFLTMG-EQN----- 211
QY 214 PPSPPSGPETAEEAALFSECRSPQAQGLRVLWLVTT-AVFFLFLCLSLYGLIGR 272
Db 212 -----RSADQHQAGGLVCTPT-IHTATVKVIVQNTFMSPFPMVVISLVNTIIAN 261
QY 273 ELWSSRRPL--RGAAS-----GREGRHQRTVRLVVLAVFLICWLPFHV 316
Db 262 KLTVMVQAEOGQVCTVGEHSTFSMAIEPGRVQALRHGVRLRAVVIAPVVCWLPFHV 321
QY 317 GRIIYINTDSR---MMY-FSQYFNIVALQFYLASINPILYNLISKYR-----AA 365
Db 322 RRLMFCVISDEQWTFELYDFHYFMYVTNALFYVSSTINILYNLVSANFRHIFLATLAC 381
QY 366 AFKLLARKSRPGRFHSRDT 386
Db 382 LCPVWRRRRKRP-AFSRKADS 401

RESULT 9

NTR2 HUMAN STANDARD; PRT; 410 AA.
AC 095665; Q8TBH6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Neurotensin receptor type 2 (NTR-2) (Levocabastine-sensitive
DE Neurotensin receptor) (NTR2 receptor).
GN NTR2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=99066919; PubMed=9851594;
RA Vita N., Oury-Donat P., Chalou P., Guillemot M., Kaghad M., Bachy A.,
Thurnissen S., Garcia S., Poinot-Chazel C., Casellas P., Keane P.,
Le Fur G., Maffrand J.-P., Shoubrie P., Caput D., Ferrara P.;
RT "Neurotensin is an antagonist of the human neurotensin NTR2 receptor
RL expressed in Chinese hamster ovary cells.";
RN Eur. J. Pharmacol. 360:265-272(1998).
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan B., Moore T., Max S.I., Wang J., Heish F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abranson R.D., Mullaby S.J.,
Boak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -I- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUROTENSIN. IT IS
CC ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
CC CALCIUM SECOND MESSENGER SYSTEM.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
CC -I- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC HIGHEST TO TACHYKININS RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; Y10148; CAAY1233.1; -.
DR EMBL; BC022501; AAH22501.1; -.
DR Genew; HGNC:8040; NTR2.
DR MIM; 605538; -.
DR GO; GO:0005887; C: integral to plasma membrane; TAS.
DR GO; GO:0007600; P: sensory perception; TAS.
DR InterPro; IPR000276; GPCR Rhodopsin.
DR InterPro; IPR003986; NTR2 rec.
DR InterPro; IPR003984; NTR2 rec.
DR Pfam; PF00001; 7tm_1_1.
DR PRINTS; PR00237; GPCR RHODOPSIN.
DR PRINTS; PR01479; NEUROTENSINR.
DR PRINTS; PR01481; NEUROTENSINR.
DR PROSITE; PS00237; G PROTEIN RECEPTOR F1_1; 1.
DR PROSITE; PS0262; G PROTEIN RECEPTOR F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Lipoprotein; Palmitate.
FT DOMAIN 1 32
FT TRANSMEM 33 55 1 (POTENTIAL).
FT DOMAIN 56 64 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 65 87 2 (POTENTIAL).
FT DOMAIN 88 109 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 110 131 3 (POTENTIAL).
FT DOMAIN 132 154 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 155 176 4 (POTENTIAL).
FT DOMAIN 177 217 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 218 237 5 (POTENTIAL).
FT DOMAIN 238 297 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 298 318 6 (POTENTIAL).
FT DOMAIN 319 337 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 338 358 7 (POTENTIAL).
FT DOMAIN 359 410 CYTOPLASMIC (POTENTIAL).
FT DISULFID 108 194 BY SIMILARITY.
FT LIPID 377 377 PALMITATE (POTENTIAL).
FT CONFLICT 168 168 L -> M (IN REF. 2).
FT CONFLICT 367 367 K -> R (IN REF. 2).
SQ SEQUENCE 410 AA; 45413 MW; 8C3ADA22BE15FD66 CRC64;
Query Match 19.2%; Score 413; DB 1; Length 410;
Best Local Similarity 31.4%; Pred. No. 8.3e-21;
Matches 130; Conservative 60; Mismatches 134; Indels 90; Gaps 16;
QY 17 PMPALPP---CDERCSPFPFGALVPTAVCLFVGVSGNVVTMLI-----GR 65
Db 8 PPRSPNPGLSLDARLGVDTRLWAKVLTALVALIWAAGAGNALSVVHVKARAGR 67
QY 66 YRDMRTTNLYLGSMASVLDLILL-GLPDLIR-LWRSRPVWVGPCLLRLSLVVGECY 123
Db 68 LRH-----HVLSTALAGLLLLVGPVELYSFVFWFVPGDGLGCGYYFHELCA 120
QY 124 ATLLHMTALSVRYLAICRPLARVLVTRRRVRLIAVLWALLSAGPFLVGV---- 179
Db 121 ATVLVAGLSAERCLAVCOPLARSLTTRTFLWVLSWAASLGLALPNAVINGQKHEL 180
QY 180 -----EODPGISVVPGLNGTARIASS-----PLASSPPI-----WLS 211
Db 181 ETADGSEPEASRVCTVL--VSR TALQVFQVNVVLSFVLPLALTAFNGVTVSHLLALCS 238

CC ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
CC CALCIUM SECOND MESSENGER SYSTEM.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: ABUNDANT IN CORTEX AND HYPOTHALAMUS, AND LOWER
CC LEVELS SEEN IN THE HEART AND INTESTINE.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED MAXIMALLY IN 7-DAY-OLD BRAIN AND
CC EXPRESSION DECREASES PROGRESSIVELY UNTIL ADULTHOOD (35-DAY-OLD
CC BRAIN).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC HIGHEST TO TACHYKININS RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X97121; CAA65787.1; -;
CC PIR; S68822; S68822.
CC InterPro; IPR000276; GPCR_Rhodopsn.
CC InterPro; IPR003986; NT2_Rec.
CC InterPro; IPR003984; NT_Rec.
CC Pfam; PF00001; 7tm.1; 1.
CC PRINTS; PR00237; GPCRHHODPSN.
CC PRINTS; PR01479; NEUROTENSINR.
CC PRINTS; PR01481; NEUROTENSIN2.
CC PROSITE; PS00237; G-PROTEIN RECEPTOR_F1_1; 1.
CC PROSITE; PS00262; G-PROTEIN RECEPTOR_F1_2; 1.
CC G-protein coupled receptor; Transmembrane; Lipoprotein; Palmitate.
CC DOMAIN 1 32
CC TRANSMEM 33 55
CC DOMAIN 56 64
CC TRANSMEM 65 87
CC DOMAIN 88 109
CC TRANSMEM 110 131
CC DOMAIN 132 154
CC TRANSMEM 155 176
CC DOMAIN 177 216
CC TRANSMEM 217 237
CC DOMAIN 238 297
CC TRANSMEM 298 317
CC DOMAIN 319 338
CC TRANSMEM 338 358
CC DOMAIN 359 416
CC DISULFID 108 194
CC LIPID 377 377
CC PALMITATE (POTENTIAL).
CC SEQUENCE 416 AA; 46265 MW; 127FC5F5CB6FE208 CRC64;
Query Match 18.68; Score 401; DB 1; Length 416;
Best Local Similarity 29.48; Pred. No. 5.3e-20;
Matches 122; Conservative 68; Mismatches 145; Indels 80; Gaps 13;
QY 18 PWPALP-----PDERRCSPFPGALVPYAVCLCLFVVGSGNVVTVMLIGRYDMRT 71
DB 6 PWPSPSPAGLSLEARLGVDTLWAKVLTALYSILFAFGTAGNALS VHVVLKARAGRP 65
QY 72 -TNNYLGSMVSDLLILL-GLPFDLYR-LWRGRPVWFGPLCLRLSLYVGCGTYATLLH 128
DB 66 GLRLYHVLSSALLSALLLVSMPEMLNFVWNSHYWVFGDLGCRGYFVRELCAATVLS 125
QY 129 MTALSRYEIAICRPLRARVLVTRRRVIALIAVMAVALLSAGPFLFLVGV----- 179
DB 126 VASLSARCLAVCOPLRRLTLPTRRRLSLVWASGLALPMAVIMQKHEVESADG 185
QY 180 EQDPSGVVPLNGTARIA-----SSPLASPPPLWLSRAPPPSPGPETAEEA 228
DB 186 EPEPASVCTVLSRATLQVFIQVNLVSPALPLATFL-----NGITVNHLM 234
QY 229 ALPSRECRPSPAQLGALRVMWVTTAYFPLCLSLYGLIGRELWSSRRPLRGPASG 288
DB 235 ALYS-QVPSASQVSSIPSRLLELSE-----EGLLGLFITWRKTLSLGVQASLV 281

QY 289 RE-----RGHRTVRLVLLVVLAFIICWLPFHVGRIIYINTED-----SRMMYFSQYFNI 338
DB 282 RHKDAQSRSLQSAQVLRVAIVAVVYICWLPYHARRLMYCYIDDGWTNELYDFYHYFM 341
QY 339 VALQLPYLGASINPIYLNISKYRAAFKLLARKSRPRGFRHSRDRDTAGEVAGD 393
DB 342 VTNTLFVSSAVTPILYNVSSFR-----KLFL-----ESLSLQGE 379
RESULT 12
SSRS_MOUSE STANDARD; PRT; 362 AA.
ID SSRS_MOUSE
AC O08858; O08998;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 13-SEP-2003 (Rel. 42, Last annotation update)
DE Somatostatin receptor type 5 (SS5R).
GN SS5R OR SMSTR5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ; TISSUE=Liver;
RX MEDLINE=97444289; PubMed=9300821;
RA Lublin A.L., Diehl N.L., Hochgeschwender U.;
RT "Isolation and characterization of the gene encoding the type 5 mouse
RT (Mus musculus) somatostatin receptor (msst5).";
RL Gene 195:63-66 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ;
RA Moldovan S., Demayo F., Brunicardi P.C.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Liver;
RA Gordon D.F., Woodmansee W.W., Wood W.M., Knauf H., James R.A.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ;
RA Baumeister H., Roosterman D., Schafer J., Kreuzer O., Meyerhof W.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: RECEPTOR FOR SOMATOSTATIN-28. THE ACTIVITY OF THIS
CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH INHIBIT ADENYLYL CYCLASE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U82697; AAC53353.1; -;
DB EMBL; AF004740; AAB61418.1; -;
DB EMBL; AF030441; AAB86492.1; -;
DB EMBL; AF035777; AAB88302.1; ALT_INIT.
DB HSP; P02699; 1F88.
DB MGD; MGI:894282; Sstr5.
DB InterPro; IPR000276; GPCR_Rhodopsn.
DB Pfam; PF00001; 7tm.1; 1.
DB PRINTS; PR00237; GPCRHHODPSN.
DB PROSITE; PS00237; G-PROTEIN RECEPTOR_F1_1; 1.
DB PROSITE; PS00262; G-PROTEIN RECEPTOR_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Multigene family; Lipoprotein; Palmitate.
FT DOMAIN 1 35 EXTRACELLULAR (POTENTIAL).
FT

"Molecular cloning and expression of a pituitary somatostatin receptor with preferential affinity for somatostatins-28." Mol. Pharmacol. 42:939-946(1992).

[2]

REVISIONS TO C-TERMINUS.

TISSUE= Pituitary;

MEDLINE=94195267; PubMed=7908405;

R Panetta R., Greenwood M.T., Warzynska A., Demchyshyn L.L., Day R.,

RA Niznik H.B., Szikant C.B., Patel Y.C.;

"Molecular cloning, functional characterization, and chromosomal localization of a human somatostatin receptor (somatostatin receptor type 5) with preferential affinity for somatostatin-28." Mol. Pharmacol. 45:417-427(1994).

-1- FUNCTION: RECEPTOR FOR SOMATOSTATIN-28. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH INHIBIT ADENYLYL CYCLASE.

-1- SUBCELLULAR LOCATION: Integral membrane protein.

-1- TISSUE SPECIFICITY: PROMINENT IN THE PITUITARY AND SMALL INTESTINE. LOW LEVELS IN ISLETS AND SPLEEN. NOT DETECTED IN KIDNEY, PANCREAS, CEREBELLUM, OR CORTEX.

-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL; L04535; AAA17029.1; -
EMBL; U01152; AAC09011.1; -
EMBL; X74828; CAA52825.1; -
PIR; I57940; I57940.
HSP; P02693; I57940.
InterPro; IPR000276; GPCR_Rhodopsin.
Pfam; PF00001; 7tm_1; 1.
PRINTS; PR00237; GPCRRHODOPSIN.
PROSITE; PS00237; G_PROTEIN_RECP_F1_1;
PROSITE; PS0262; G_PROTEIN_RECP_F2_1;
G-protein coupled receptor; Transmembrane; Glycoprotein;
Multigene family; Lipoprotein; Palmitate.
DOMAIN 1 35 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 36 63 1 (POTENTIAL).
FT DOMAIN 64 73 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 74 99 2 (POTENTIAL).
FT DOMAIN 100 111 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 112 133 3 (POTENTIAL).
FT DOMAIN 134 155 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 156 176 4 (POTENTIAL).
FT DOMAIN 177 196 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 197 221 5 (POTENTIAL).
FT DOMAIN 222 247 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 248 273 6 (POTENTIAL).
FT DOMAIN 274 283 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 284 308 7 (POTENTIAL).
FT DOMAIN 309 363 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 13 13 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 23 23 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC...) (POTENTIAL).
FT DISULFID 110 185 BY SIMILARITY.
FT LPID 320 320 PALMITATE (POTENTIAL).
SQ SEQUENCE 363 AA; 39971 MW; 4BD4512960613B4A CRC64;

Query Match 16.7%; Score 359; DB 1; Length 363;
Best Local Similarity 27.3%; Pred.No.3e-17;
Matches 110; Conservative 60; Mismatches 149; Indels 84; Gaps 12;

Qy 5 WNGSDPEGARBPWPALPCDERRCSPPRLGA---LVPTAVTCVLCDFVVGVSGSVNVTVVM 61
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| : : : : :
12 WNASASSGNHN--WSLVG-----SASPMGRADVLPVLVLC--TVGLSGNTLVIV 60
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| : : : : :
Qy 62 LIQRYDRMTTTNLYLGSMVASVDLLILGLFPDLRYLRMRSPWVPGLLCRLSLYYVGEC 121
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| : : : : :

FT	TRANSMEM	284	308	7 (POTENTIAL).	
FT	DOMAIN	309	363	CYTOPLASMIC (POTENTIAL).	
FT	CARBOHYD	13	13	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	23	23	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	186	186	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	DISULFID	110	185	BY SIMILARITY.	
FT	LIPID	320	320	PALMITATE (POTENTIAL).	
SQ	SEQUENCE	363 AA;	35971 MM;	4BD4512960613B4A	CRC64;
	Query Match		16.7%;	Score 359;	DB 1; Length 363;
	Best Local Similarity	27.3%;	Pred. No. 3e-17;		
	Matches 110;	Conservative	60;	Mismatches 149;	Indels 84; Gaps 12;
Qy	5	WNGSDGEGAREPPWALP	PCDERRCSPPPLGA---	LPVVTAVCLCLFVVGVGSNVVTVM	61
		WNGS			
Db	12	WNASAASSGNHN--WSLVG	-----SASPMGARAVLPVLYLLVC-	TVGLSGNTLVIY	60
		WNAS			
Qy	62	LIGRYDRMRTTNNLYLG	SMAVSDLLILGLGFPDLYRLWRSRPWP	FPGLLCRLSLYVGEGC	121
		LIGRY			

Db 61 VVLRHAKMKTVTNVIILNLAVADVFLMLGLPFLATQNAVVSYPFGSFLCRLVMTLDGIN 120
 QY 122 TYATLLHMTALSVERYLAICRPLARLVTRRRVRLIALVLAVALLSAGPFLFLVGVEQ 181
 Db 121 QFTSIFCLMWSVDRIYLVVHPILRSARWRPRVAKASAAVWFSLMSLPLLVFADVQE 180
 QY 182 DPGISVVVGLNGTARIASSPLASSPPLWLRAPPPSPFGPETAFAAALFSRECRPSPAQ 241
 Db 181 G-----WGICNLSS-----W-----PEPVG 194
 QY 242 LGLARVWLWVTAYFFFLPFLCLILYGLI-----GRELWSSRRRLRGPAAASGRBRGHR 294
 Db 195 LWGAFTTYSVLGFFGLFVLICVLLVVKVKAAGMRVGSRR-----RRRSEP 244
 QY 295 QTVRVLLAVVLAFLICWLPVHGRIIYN-TEDSRMWFQYNSRIVALQFLYLSASINPI 353
 Db 245 KVTSMVVVVVLFVGCWLPFFIVNVMLATLPEEPTSAGLYFFVVVLS--YANSCANPL 302
 QY 354 LYNLSIKKYRAAFAKLLARKSRPRGHRSDTAGVAGDTGG 396
 Db 303 LYGLSLDNFQSRFRKVLCLR----RGYGMEDADAIEPRPKSG 341

RESULT 14
 SSR3 HUMAN
 ID SSR3 HUMAN STANDARD; PRT; 418 AA.
 AC P32745;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Somatostatin receptor type 3 (SSR3) (SSR-28).
 GN SSR3.
 OS Homo sapiens (human).
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93149123; PubMed=1337145;
 RA Yamada Y., Reisine T., Law S.F., Ihara Y., Kubota A., Kagimoto S.,
 RA Seino M., Seino Y., Bell G.I., Seino S.;
 RT "Somatostatin receptors, an expanding gene family: cloning and
 RT functional characterization of human SSR3, a protein coupled to
 RT adenylyl cyclase.";
 RL Mol. Endocrinol. 6:2136-2142(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93238970; PubMed=8097479;
 RA Corness J.D., Demchyshyn L.L., Seeman P., van Tol H.H.M.,
 RA Srikant C.B., Kent G., Patel Y.C., Niznik H.B.;
 RT "A human somatostatin receptor (SSR3), located on chromosome 22,
 RT displays preferential affinity for somatostatin-14 like peptides.";
 RL FEBS Lett. 321:279-284 (1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20057165; PubMed=10591208;
 RA Dunham I., Hunt A.R., Collins J.E., Bruskiewich R., Beare D.M.,
 RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
 RA Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
 RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
 RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
 RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
 RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
 RA Dhani P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,
 RA Evans K.B., Fey J.M., Fleming K., French L., Garner A.A.,
 RA Gilbert J.G.R., Goward M.B., Grafham D.V., Griffiths M.N.D., Hall C.,
 RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,
 RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,
 RA Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,
 RA Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,
 RA Mcclay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,
 RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
 RA Phillips S.H., Plumb R.W., Ramsey H., Ramsey Y., Rogers L., Ross M.T.,

RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
 RA Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M.,
 RA Vaudin M., Walli M., Wallis J.M., Whiteley M.N., Willey D.L.,
 RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
 RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
 RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
 RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuayama S.,
 RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
 RA Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,
 RA Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,
 RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,
 RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
 RA Zhan M., Zhang G., Chisoe S., Murray J., Miller N., Minx P.,
 RA Fulten R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,
 RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
 RA Hinds K., Kemp K., Latreille P., Layman D., Ozersky P., Rohlfing T.,
 RA Scheet P., Walker C., Wamsley A., Wohldmann P., Pepin K., Nelson J.,
 RA Korf I., Bedell J.A., Hillier L., Hardis E., Waterston R., Wilson R.,
 RA Emanuel B.S., Shaikh T., Kurahashi H., Saitta S., Budarf M.L.,
 RA McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E., Edelman L.,
 RA Kim U.J., Shizuya H., Simon M.I., Dumanski J.P., Peyrard M., Kedra D.,
 RA Seroussi E., Franssen I., Tapia I., Bruder C.E., O'Brien K.P., Lane L.,
 RA Wilkison P., Bodenteich A., Hartman K., Hu X., Khan A.S., Lane L.,
 RA Tilahun Y., Wright H.;
 RT "The DNA sequence of human chromosome 22.";
 RL Nature 402:489-495(1999).
 CC -1- FUNCTION: RECEPTOR FOR SOMATOSTATINS-14 AND -28. THIS RECEPTOR IS
 CC COUPLED VIA PERTUSSIS TOXIN SENSITIVE G PROTEINS TO INHIBITION OF
 CC ADENYLYL CYCLASE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: BRAIN, PITUITARY AND PANCREAS.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M96738; AAA60592.1; -;
 CC EMBL: Z82188; CAB45263.1; -;
 CC PIR: A46226; A46226.
 CC HSSP: P34996; LDDD.
 CC Genew; HGNC:11332; SSTR3.
 CC MIM; 182453; -;
 CC GO; GO:0005887; C:integral to plasma membrane; TAS.
 CC GO; GO:0004994; F:somatostatin receptor activity; TAS.
 CC GO; GO:0007267; P:cell-cell signaling; TAS.
 CC GO; GO:0007187; P:G-protein signaling; coupled to cyclic nucl. . . ; TAS.
 CC GO; GO:0008628; P:induction of apoptosis by hormones; TAS.
 CC GO; GO:0008285; P:negative regulation of cell proliferation; TAS.
 CC InterPro; IPR000276; GPCR_Rhodpsn.
 CC Pfam; PF00001; 7tm 1; 1.
 CC PRINTS; PR00237; GPCR_Rhodpsn.
 CC PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
 CC PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
 CC G-protein coupled receptor; Transmembrane; Glycoprotein;
 CC Multigene family; Polymorphism.
 CC DOMAIN 1 43 EXTRACELLULAR (POTENTIAL).
 CC TRANSMEM 44 69
 CC 1 (POTENTIAL).
 CC CYTOPLASMIC (POTENTIAL).
 CC DOMAIN 70 79
 CC 2 (POTENTIAL).
 CC TRANSMEM 80 101
 CC 3 (POTENTIAL).
 CC DOMAIN 102 116
 CC 4 (POTENTIAL).
 CC TRANSMEM 117 138
 CC 5 (POTENTIAL).
 CC DOMAIN 139 161
 CC 6 (POTENTIAL).
 CC TRANSMEM 162 181
 CC 7 (POTENTIAL).
 CC DOMAIN 182 205
 CC 8 (POTENTIAL).
 CC TRANSMEM 206 231
 CC 9 (POTENTIAL).
 CC DOMAIN 232 257
 CC 10 (POTENTIAL).
 CC TRANSMEM 258 279
 CC 11 (POTENTIAL).
 CC DOMAIN 280 293
 CC 12 (POTENTIAL).

FT TRANSMEM 294 316 7 (POTENTIAL).
 FT DOMAIN 317 418 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 17 17 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 30 30 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT DISULFID 116 191 BY SIMILARITY.
 FT DOMAIN 346 360 GLU-RICH (ACIDIC).
 FT VARIANT 411 411 S->T (IN dSNP:229568).
 FT SEQUENCE 418 AA; 45847 MW; 1227095F801190C4 CRC64;
 Query Match 16.6%; Score 357.5; DB 1; Length 418;
 Best Local Similarity 26.9%; Pred. No. 4.2e-17;
 Matches 101; Conservative 63; Mismatches 146; Indels 65; Gaps 10;
 QY 11 PEGAREPPWALPCDERRCSPPFL-----GALVPVTAACVLCFLVGVGSGNVVTMLIGR 65
 DB 15 PENA-SSAWPPDATTGNVSGSPAGLAVSGVLIPLVYLVC--VGLLGNLSLVIYVVL 71
 QY 66 YRDMRTTNLYLGSMAVSDLLILGLPDLRLWRSPWVFGPLLCRLSLYVGEQTYAT 125
 DB 72 HTASPSVTNYIINLALADELMGLGFP-LAAQNALSYWPFGLMCLVMAVDGINQFTS 130
 QY 126 LLHMTALSVRYLAICRPLARVLVTRRRVRLAVLMAVALLSAGPFLVGVQDQPGI 185
 DB 131 IFCLTVMSVDRLAVHPTARSARWTPARTVAAVMAVAVVVLVPPVVFSGV----- 184
 QY 186 SVVPGGLNGTARIASSPLASSPPLWLSRAPPPSPGPETAEEAALFSRECR---PSPAOL 242
 DB 185 -----PRGMST-----CHMQWPEPAAA 201
 QY 243 GALTVMVWVTAFFLPFLCLSLYGLIGRELWSSRLRGPAAAGRGHROTQVRLV 302
 DB 202 WRAGFIITTAALGFGPLLVICLYLLVTVKRSAGRRVWAPSQRRRRSERRVTRMVA 261
 QY 303 VLAFLICMLPEHVGRIIYINTE-DSRMVFSQVENVALQLEVLASINPILNLSKK 361
 DB 262 VVALFVLCWMPFFVNLVNVVCPLEPAPFAFFLYFLVALP--YANSCANPILYGFUSYR 319
 QY 362 YRAAFKLLARKSR 376
 DB 320 FK-QGFRVLLRPSR 333

RESULT 15

GP39 HUMAN STANDARD; PRT; 453 AA.
 AC O43194;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Putative G protein-coupled receptor GPR39.
 GN GPR39.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RC MEDLINE=98110578; PubMed=9441746;
 RA McKee K.K., Tan C.P., Palyha O.C., Liu J., Feighner S.D.,
 RA Hreniuk D.L., Smith R.G., Howard A.D., van der Ploeg L.H.T.;
 RT "Cloning and characterization of two human G protein-coupled receptor
 genes (GPR38 and GPR39) related to the growth hormone secretagogue
 and neurotensin receptors.";
 RT Genomics 46:426-434(1997).
 RL CC
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.emb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; AF034633; AAC26082.1; -
 DR Genew; HGNC:4496; GPR39.
 DR MIM; 602886; -
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.
 DR GO; GO:0007186; P:G-protein coupled receptor protein signal. . .; TAS.
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm 1; 1.
 DR PRINTS; PR00237; GPCRHOPOPSN.
 DR PROSITE; PS00237; G-PROTEIN-RECEP_F1_1; 1.
 DR PROSITE; PS00262; G-PROTEIN-RECEP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein.
 FT DOMAIN 1 25 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 26 46 1 (POTENTIAL).
 FT DOMAIN 47 71 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 72 92 2 (POTENTIAL).
 FT DOMAIN 93 109 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 110 131 3 (POTENTIAL).
 FT DOMAIN 132 154 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 155 175 4 (POTENTIAL).
 FT DOMAIN 176 223 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 224 243 5 (POTENTIAL).
 FT DOMAIN 244 280 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 281 301 6 (POTENTIAL).
 FT DOMAIN 302 321 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 322 347 7 (POTENTIAL).
 FT DOMAIN 348 428 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 108 210 BY SIMILARITY.
 FT CARBOHYD 192 192 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 206 206 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 212 212 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 453 AA; 51328 MW; 8E3A233420D9021E CRC64;

Query Match 16.5%; Score 355.5; DB 1; Length 453;

Best Local Similarity 28.5%; Pred. No. 6.2e-17;

Matches 111; Conservative 65; Mismatches 169; Indels 45; Gaps 13;

QY 20 PALPPCDERRCS-----PPPLGALVPVT--AVCLCLFVGVGSGNVTV---MLIGR 65
 DB 4 PSLPGSD---CSQIIDHSHVPEFVATWIKITILVLIIFVMLGLGNSATIRVTVLQK 60

QY 66 YRDMRTTNLYLGSMAVSDLLI-LLGLPFDLYR-LWRSRWVFGPLLCRLSLYVGEQTY 123
 DB 61 KGYLQKEVTDHVMVSLACSDILVELIGNPMEFYIINPLTTSYTLCKLHTELPACSY 120

QY 124 ATLLHMTALSVRYLAICRPLARVLVTRRRVRLAVLMAVALLSAGPFLVGVQDQ 183
 DB 121 ATLLHVLTLSEFYIAICHPRYKAVSGPCQVKLLIGFVWVTSALVALPLLPANGTEY-P 179

QY 184 GISVVGGLNGTARIASSPLASSPPLWLSRAPPPSPGPETAEEAALFSRECPSPAQ-- 241
 DB 180 LVN-VPSHRLGTLCSSTRHEQ-----PETSMSICTNLSSRWTVFQSS 223

QY 242 -LGALRVMLWVTTAYFPFLPCLSLYGLIGRELWSSRRP--LRGPAASGRGRHROTQV 298
 DB 224 IFGAFVYLVLLSVAFMCMNMVQVLMKQSGSLAGTRPPQLRKSESESRATRRITII 283

QY 299 VLVVAVLAFIICWLPFHVGRIIYI--NTEDSRMMYFSQYFNIVALQ--LPYLSASINPIL 354
 DB 284 FLRLIVTVLAVCMVNPQIRIRMAAAKPKHDWTSYFRAYMILLPFPSTFYLSSVINPLL 343

QY 355 YNLSIKKYRAAAPKLLARKSRGRGRHRSR 384
 DB 344 YTVSSQOFRFRVQVQLCCRLSLOHANHEKR 373

Search completed: January 1, 2004, 06:23:01

Job time : 16.4561 secs

THIS PAGE BLANK (USPTO)